



Align seg 1/1 to: AAB47096 from: 1 to: 120

[illegible]

seq\_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT: AAB47097

```
seq_documentation_block:
```

ID AAB47097 standard; Protein; 174 AA

AC	AA047097;
XX	
DT	16-MAY-2001 (first entry).
XX	
DE	Korean mistletoe lectin #2.
XX	
XX	isoform; A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
KM	KM: tumour; KM-110; KM-C; KMAB; KM-IT0; KM-ITL;
KW	heparin binding protein.
XX	
OS	Viscum album coloratum.
XX	
PM	EP1074560-A2.
XX	
PD	07-FEB-2001.
XX	
XX	27-JUL-2000; 2000EP-0402168.
PE	
XX	27-JUL-1999; 99KR-0030638.
PR	
PA	(MIST-) MISTLE BIOTECH CO LTD.
PI	
PI	Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
PI	Kang T, Park C;
PI	
DR	WPI: 2001-171044/18.
DR	N-PSDB: AHC85479.
XX	
PT	Novel lectin proteins isolated from Korean mistletoe useful for
XX	enhancing immunity and effectuating anti-tumoral activity -
PS	
PS	Claim 33: Page 34; 62pp; English.
XX	
XX	The sequences given in AA047096-97 are lectins isolated from Korean

CC mistletoe. Korean mistletoe lectins (KML) are useful for enhancing  
CC immunity and for treating tumours. The KMLs are isolated from a  
CC protein fraction derived from the leaves, stems and fruits of Korean  
CC mistletoe, which is designated KM-10. One of the isolates, KML-C was  
CC shown to be effective against colon 26-M3.1 carcinoma and LS178Y-M25  
CC lymphoma.

CC Sequence 174 Aa.

CC 2X

CC 20

```
alignment_scores:      408.50      1222
                    Quality:
                    Ratio: 4.045      1
Percent Similarity:    82.787      Percent Identity: 69.672
```

alignment\_block:  
US-09-627-165B-13 x AAB47097

Align seg 1/1 to: AAB47097 from: 1 to: 174

1	GCACATTTCAATCCCATCCTGAGAGGCTCCGCCGCAATTAACAGTGG	50
53	AlaAagPheAsnProIlePheTrpArgAlaArgInTrpIleAsnSerCyl	100
55	GGAGTCNTCTCCACCAACATGTGATCATCTCTGAGCTGAGACAGATTGG	60
69	YglUuSerPheCecPheGAspMetTrpMetLleAluGluSerGluTrpP	86
101	GTGCGACACACCAACAGCTCAGACAGTCCAAAGATATGGCATTTTAAATAC	150
86	YglGlnGlnSerTrfGlnAlaGlnGlnSerTrhAspLysAlaPheAsnAsn	100
151	CAATATAGATGTGCATTTTCCGCCGTAACTCTGTGACGNTGACCAATGT	200
103	ProPheArgLysGluYlaSerTrhArgLysMetPheValTrhLeuSerAsn	118
201	TGGCAGATGATCATCGAGTGGCATGATGTTGTCACAAAGAGAGTGC	250
119	ArgAspAlaIleAlaSerLeuAlaIleMetLeuPheValCysArgAspA	136
251	GGCCATTTCTCTCTCTGCACACACCTTGGCCGCTGCTCTTAAGTCCGTC	300
136	TrpProSerSerSerAspValAlaTrpTrpProLeuAlaIleAspGlyAla	157
301	GTGAGTCT	344
153	LeuGlnAsnSerCylYalaValAlaAspAlaTrhCysTrhAlaSerGluTr	169
342	CACCGTGGGATCGTA	357
169	othrValArgIleVal	174

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:AAW10021

```
seq_documentation_block:
```

ID AAW10021 standard; Protein; 564 AA

AC	AAW10021;	
DT	18-DEC-1997 (first entry)	
DX	Prepro mistletoe lectin.	
DX	Mistletoe lectin: cytotoxic; A chain; B chain; dimer	
KW		
XX	Viscum album.	
OS		
XX	EF751221-AL.	
XX		
PD	02-JAN-1997.	
XX		
PF	26-JUN-1995.	95EP-0109949
XX		

PR 26-JUN-1995; 95EP-0109949.  
XX (MADU ) MADNUS KOELN AG.  
XX Baur A, Eck J, Lentzen H, Zinke H;  
XX MPI: 1997-054678/06.  
DR N-PSDB: AAT70473.  
XX  
XX Nucleic acid encoding pre:pro form of mistletoe lectin - for  
PT therapeutic or diagnostic use  
XX  
XX Claim 12: Fig 4c; 30pp; German.  
XX  
XX Mistletoe lectin is a cytotoxic agent that has been used for tumour  
CC therapy. It can be used in immunotoxins and medicaments. Nucleic  
CC acid fragments can be used in diagnostic methods. Mistletoe lectin (  
CC AAT70473) comprises an A chain (AAT70474) and a B chain (AAT70475).  
XX  
SQ Sequence 564 AA;

alignment\_scores:  
Quality: 387.00 Length: 119  
Ratio: 3.990 Gaps: 1  
Percent Similarity: 81.513 Percent Identity: 67.227

alignment\_block:  
US-09-627-165B-13 x AAM10021 ..

Align seg 1/1 to: AAM10021 from: 1 to: 564

1 GCCAGATTCATCCCATGNTGGAGGCTTCGCCGCAATTACAGTGG 50  
|||||  
200 AAtAgPheAsnProIleuTrpArgAlaAgInTyRileAsnSergI 216  
51 GGAGCNCNTCCACCAACATGTACATGCTGCAGCTGGAGAGCGAGTGGG 100  
|||||  
216 yAlaserPheleuProAspAllyrMetleuGluLeuGlnThrsertIpg 233  
101 GTGCACAAATCCACCAAGTCCAGAGTCCAGAGATGGAATTTAAATACC 150  
|||||  
233 yGInGInSerThrGInValGInHIsSerThrAspGlyValPheAsnAsn 249  
151 CAATATAGATTCAGATTTCCGCCGGTAATCTTGGAGAGTACAGATGT 200  
|||||  
250 ProIleArgLeuAlaIleProProGlyAsnPheValThrLeuThrAsnVa 266  
201 TGGGAGCGTATCTCCAGCTTGGCGATCATGTTGTCGAATGCGAGTGGC 250  
|||||  
266 lArgAspValIleAlaSerleuAlaIleMetleuPheValIcysGlyGln 283  
251 GGCATTTCTCTCTGACACCCTTCGCCGCTGCTCTTAAGTCGCTC 300  
|||||  
283 rPProSerSerSergIuValArgTyTrpProLeuValIleArgProVal 299  
301 GTGGATGGCGCAAGATGTACATGCTACATNTTTCGAACCCAGCGTGGC 350  
|||||  
300 Ile.....AlaAspAspValThrCysSerAlaSerGluProThrValAr 314  
351 CATCGTA 357  
314 gIleVal 316

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:AAW90127

seq\_documentation\_block:

ID AAW90127 standard; Protein: 564 AA.  
XX  
XX AAW90127;  
AC  
XX  
XX 30-APR-1999 (first entry)  
XX

DE Mistletoe lectin prepro-protein.  
XX  
XX MW: mistletoe; lectin; ML: transgenic plant; glycosylation;  
KW dimer; immunotoxin; large-scale production; diagnosis; therapeutic;  
KW cancer.  
XX  
XX Viscum album.  
XX  
XX EP84388-A1.  
XX  
XX 16-DEC-1998.  
XX  
XX 26-JUN-1995; 95EP-0109949.  
XX  
XX 26-JUN-1995; 95EP-0109949.  
PR 26-JUN-1995; 98EP-0105660.  
XX  
XX (MADU ) MADNUS KOELN AG.  
XX  
XX Baur A, Eck J, Lentzen H, Zinke H;  
XX MPI: 1999-026582/03.  
DR N-PSDB: AAV74182.  
XX  
XX New transgenic plant expressing mistletoe lectin - useful for  
PT producing recombinant lectin in e.g. cancer diagnosis and therapy  
XX  
XX Claim 1a: Fig 4c; 30pp; German.  
XX  
XX This invention describes a novel transgenic plant transformed with a  
CC vector capable of encoding a mistletoe (Viscum album) lectin  
CC preprotein or a biologically active fragment. The specification  
CC also describes a polypeptide produced by a plant where the polypeptide  
CC exhibits at least one enzymatic modification other than the glycosylation  
CC that occurs in viscum album or the polypeptide is a fusion protein, a  
CC mistletoe lectin polypeptide dimer and an immunotoxin comprising the  
CC polypeptide or the polypeptide dimer. The plants are used for large-scale  
CC production of mistletoe lectin for diagnostic or therapeutic purposes  
CC (e.g. in cancer therapy). This sequence represents the mistletoe lectin  
CC used in the method of the invention.  
XX  
SQ Sequence 564 AA;

alignment\_scores:  
Quality: 387.00 Length: 119  
Ratio: 3.990 Gaps: 1  
Percent Similarity: 81.513 Percent Identity: 67.227

alignment\_block:  
US-09-627-165B-13 x AAW90127 ..

Align seg 1/1 to: AAW90127 from: 1 to: 564

1 GCCAGATTCATCCCATGNTGGAGGCTTCGCCGCAATTACAGTGG 50  
|||||  
200 AAtAgPheAsnProIleuTrpArgAlaAgInTyRileAsnSergI 216  
51 GGAGCNCNTCCACCAACATGTACATGCTGCAGCTGGAGAGCGAGTGGG 100  
|||||  
216 yAlaserPheleuProAspAllyrMetleuGluLeuGlnThrsertIpg 233  
101 GTGCACAAATCCACCAAGTCCAGAGTCCAGAGATGGAATTTAAATACC 150  
|||||  
233 yGInGInSerThrGInValGInHIsSerThrAspGlyValPheAsnAsn 249  
151 CAATATAGATTCAGATTTCCGCCGGTAATCTTGGAGAGTACAGATGT 200  
|||||  
250 ProIleArgLeuAlaIleProProGlyAsnPheValThrLeuThrAsnVa 266  
201 TGGGAGCGTATCTCCAGCTTGGCGATCATGTTGTCGAATGCGAGTGGC 250  
|||||  
266 lArgAspValIleAlaSerleuAlaIleMetleuPheValIcysGlyGln 283

```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AAI999.DAT:AAV25979
seq_id: AAV25979 standard; Protein; 531 AA.
seq_documentation_block:
ID      AAV25979 standard; Protein; 531 AA.
AC      AAV25979;
XX
XX      18-OCT-1999 (first entry)
DR
DE      Mistletoe lectin I protein fragment.
KA      Mistletoe; lectin; antitumour; immunostimulant; A-chain; MIA; immunity;
KW      lysozyme 26S subunit; non-cytotoxic; T-cell activation; immune response;
KW      lymphokine-producing macrophages; uncontrolled cell growth; treatment;
KW      cancer; cytotoxicity; antigen; isoform; lectin I.
XX
XX      Viscum album.
OS
XX      DE19804210-A1.
PN
XX      12-AUG-1999.
PD
XX      03-FEB-1998; 98DE-1004210.
PF
XX      03-FEB-1998; 98DE-1004210.
PR
XX      (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
PA
XX      Morris P, Stiefel T, Voelter W, Welters F;
PI
XX      MPI: 1999-445335/38.
DR
XX      N-PDB; AAQ09103.
PT
XX      Preparation of mistletoe lectins in heterologous systems,
PT      particularly for use as anticancer agents and immunostimulants
PS      Claim 7; Fig 1B; 78bp; German.
XX
XX      This invention describes a novel mistletoe lectin (I) and its fragments
XX      which have antitumour and immunostimulatory activity. The A-chain (MIA)
XX      of the mistletoe lectin binds to, and inactivates, the 26S subunit of
XX      ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX      lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX      fragments are used to treat uncontrolled cell growth (particularly of
XX      cancers) and if they lack cytotoxicity, to increase the strength of the
XX      immune response, particularly to a co-administered antigen
XX      (tumour-associated, bacterial or viral). The method allows production of
XX      mistletoe lectin, and its individual chains, in many different isoforms
XX      and on a large scale, at any time of the year. Recombinant products are
XX      free from toxins present in natural mistletoe extracts. This sequence
XX      represents a mistletoe lectin I protein fragment.
SQ      Sequence 531 AA:

Alignment_scores:          Quality: 386.00          Length: 119
                          Ratio: 3.979              Gaps: 1
Percent Similarity: 81.513 Percent Identity: 67.227

```

```

Alignment_block:
US-09-627-165B-13 x AAY25979  ..

Align seq 1/1  to: AAY25979  from: 1  to: 531

1  GCCAGATTCAATCCCATCTCTGGAGGCTTGGCCGCAATTATACAGTGG  50
   (((((((((((((((((((((((((((((((((((((((((((((((
167  AATATpHeaSnPrOllEaLseRPaLaRgAlaGdInLyLleInSeRd  183
   51  GGAATCTCTCCACCAACATGTATCTGCTGAGCTGGAGACAGATGGG  100
      |||  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
183  yALaSeRPhLeuPrOaSpVaLyLLeTLeuGdInLeGdInThSeRyRg  200
   101  GTGCAGATATCCACCCAGTTCAGACATCCAGAGATGGCATTTTAATCC  150
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
200  LyGInGInSeRThRcInLyGInLInLInSeRThInaRgLyVaLpHeaSn  216
   151  CAATTAAGATTGCAGATTCTGGCGGGTAACTTTTGTGACGATGAGCAATG  200
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
217  PrOlaErgLeuAlAlleRPrOgLyAnRPhLeAlThRleuThRAnVa  233
   201  TCCGACAGTGATTCAGACTTCGACATCATGTGTTTGAATGCAGTGGTC  250
      23  IATgAPaVaLlLeaLseRLeuAlLeLLeuLpHeVaLySgLyGuA  250
      251  GGCCATTCTCTCTCTGACACACCCCTTGCGCGCTGCTCTAAAGTCCGTC  300
          |||||  |||||  ::  |||||:::|:::|:::|:::|:::|
250  rPrOSeRSeRAsPaVaLaRgLyTrPrPrOleuVallLeaRPrOvaL  266
   301  GTGATATCGCGCCAGATGACCTGATCACTATTTTCGAAACCCAGTGGC  350
      :::  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
267  lLe.....AlaSPaSPaVaLThRcYSeRAlaSeRgLnPrOThVaLaR  281
   351  CATCGTA 357
      |||||
281  gLlEaL 283

seq_name: /SIDS/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.AAY25982
seq_documentation_block:
ID AAY25982 standard; Protein: 532 AA.
AC AAY25982;
XX
XX 18-OCT-1999 (first entry)
XX
XX
DE Mistletoe lectin I (variant) protein fragment.
XX
XX Mistletoe; lectin; antitumor; immunostimulant; A-chain; MIA; immunity;
XX KM lymphocyte 285 subunit; non-cytotoxic; T-cell activation; immune response;
XX KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX KM cancer; cytotoxicity; antigen; isoform; lectin I.
XX
XX Vascum album.
XX
XX DE19804210-AL.
XX
XX 12-AUG-1999.
XX
XX 03-FEB-1998; 98DE-1004210.
XX PF
XX PR 03-FEB-1998; 98DE-1004210.
XX
XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
XX Morris P, Stiefel T, Voelter W, Welters P;
XX
XX WPI: 1999-445335/38.
XX DR N-PSDB: AA209106.
XX
XX Preparation of mistletoe lectins in heterologous systems,
XX particularly for use as anticancer agents and immunostimulants
XX

```



PS Disclosure: Fig 4B: 78pp: German.

XX This invention describes a novel mistletoe lectin (I) and its fragments  
CC which have antitumor and immunostimulatory activity. The A-chain (MLA)  
CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of  
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and  
CC lymphokine-producing macrophages, so stimulate immunity (I) and its  
CC fragments are used to treat uncontrolled cell growth (particularly  
CC cancers) and if they lack cytotoxicity, to increase the strength of the  
CC immune response, particularly to a co-administered antigen  
CC (tumor-associated, bacterial or viral). The method allows production of  
CC mistletoe lectin, and its individual chains. In many different isoforms  
CC and on a large scale, at any time of the year. Recombinant products are  
CC free from toxins present in natural mistletoe extracts. This sequence  
CC represents a fragment of a mistletoe lectin I protein variant.

XX Sequence 532 AA;

Alignment\_scores:

Quality: 386.00 Length: 119  
Ratio: 3.979 Gaps: 1  
Percent Similarity: 81.513 Percent Identity: 67.227

Alignment\_block:

US-09-627-165B-13 x AAY25982 ..

Align seq 1/1 to: AAY25982 from: 1 to: 532

```
1 GCCGATTTCATCCATCMTGTGAGCCTTCGCCGCAATTAACAGTGG 50
|||||
167 AlaArgPheAsnProIleuTrpArgAlaArgIleuTrpIleAsnSerG 183
51 GGAATCMTTCGCACCAACATGATACATCTGACGTGAGCAGTGGG 100
| ||| |||
183 YAlSerPheLeuProAspValTyrMetLeuGluLeuGluThrSertips 200
101 GTCGCAATTCACCAAGTCACAGTCAGTCAGATGGCATTTATATAC 150
|||||
200 IYGLNGInSerThrGlnValGlnHisSerThrAspValIlePheAsn 216
151 CAATTAAGATTCGATTTCCGCCGATCTTGTGACGNTGACATGT 200
|||||
217 ProIleArgLeuAlaIleProProGlnAsnPheValThrLeuThrAsn 233
201 TCGGACGTGATCTCCAGCTTCGCGATCATGTTGTTGCAATGCAATG 250
|||||
233 IArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysGlyAla 250
251 GGCATTCCTCCTCTCGACACACCTTCGCCGCTGCTCAAGTCCGTC 300
|||||
250 rGProSerSerSerAspValArgTyrTrpProLeuValIleArgProVal 266
301 GTGATGCGCGCAACGATGTCACTGACACTTTTCCACACCCGCTGG 350
:::
267 Ile.....AlaAspAspValThrCysSerAlaSerIleProThrValAr 281
351 CATCGTA 357
|||||
281 gIleVal 283
```

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT: AAY25970

seq\_documentation\_block:

ID AAY25970 standard; protein: 533 AA.

XX AAY25970:

XX 18-OCT-1999 (first entry)

XX Mistletoe lectin protein consensus sequence 1.

XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;

KM ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;  
KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;  
KM cancer; cytotoxicity; antigen; isoform.

XX Viscum album.

```
XX Key location/Qualifiers
FH Misc-difference 15
FT /label= Asp, Glu
FT Misc-difference 63
FT /label= Gly, Gln
FT Misc-difference 66
FT /label= Ile, Val
FT Misc-difference 75
FT /label= Leu, Ala
FT Misc-difference 107
FT /label= Asp, Arg or none
FT Misc-difference 113
FT /label= Asn, Thr
FT Misc-difference 117
FT /label= Pro, Thr
FT Misc-difference 134
FT /label= Asp, Glu
FT Misc-difference 141
FT /label= Ser, Thr
FT Misc-difference 145
FT /label= Phe, Tyr
FT Misc-difference 152
FT /label= Thr, Ala
FT Misc-difference 177
FT /label= Ala, Tyr
FT Misc-difference 180
FT /label= Tyr, Asp
FT Misc-difference 185
FT /label= Ala, Glu
FT Misc-difference 191
FT /label= Val, Met
FT Misc-difference 219
FT /label= Ile, Phe
FT Misc-difference 224
FT /label= Pro, Ser
FT Misc-difference 225
FT /label= Pro, Thr
FT Misc-difference 232
FT /label= Thr, Ser
FT Misc-difference 236
FT /label= Asp, Ser
FT Misc-difference 287
FT /label= Asn, Ser
FT Misc-difference 290
FT /label= Cys, Arg
FT Misc-difference 325
FT /label= Gly, Asn
FT Misc-difference 364
FT /label= Gly, Asp
FT Misc-difference 426
FT /label= Gly, Gln
FT Misc-difference 435
FT /label= Val, Asp
FT Misc-difference 439
FT /label= Gln, Lys
FT Misc-difference 442
FT /label= Gly or none
FT Misc-difference 443
FT /label= Arg, Lys
FT Misc-difference 464
FT /label= Cys, Ser, Val
FT Misc-difference 480
FT /label= Ala, Gly
FT Misc-difference 481
FT /label= Gly, Ala
FT Misc-difference 483
FT /label= Ser, Gly
```

```

FT Misc-difference 484 /label= Gly, Ser
FT Misc-difference 493 /label= Gly, Tyr
FT Misc-difference 500 /label= Asn, Ser, Thr, Lys
FT Misc-difference 501 /label= Ser, Gly
FT Misc-difference 502 /label= Leu, Pro
FT Misc-difference 503 /label= Ala, Met
FT Misc-difference 504 /label= Met, Val
FT Misc-difference 533 /label= Pro, Phe
FT
FN DE19804210-A1.
PD 12-AUG-1999.
PE
PP 03-FEB-1998; 98DE-1004210.
PR 03-FEB-1998; 98DE-1004210.
PS (BIO5-) BIOSYN ARZNEIMITTEL GMBH.
PI Morris P, Stiefel T, Voelter W, Walters P;
PI WPI; 1999-445335/38.
PR Preparation of mistletoe lectins in heterologous systems,
PT particularly for use as anticancer agents and immunostimulants
PS Claim 1, Page 25-26; 78pp; German.
XX
XX This invention describes a novel mistletoe lectin (I) and its fragments
XX which have antitumor and immunostimulatory activity. The A-chain (KMA)
XX of the mistletoe lectin binds to, and inactivates, the 28S subunit of
XX ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX fragments are used to treat uncontrolled cell growth (particularly
XX cancers) and if they lack cytotoxicity, to increase the strength of the
XX immune response, particularly to a co-administered antigen
XX (tumor-associated, bacterial or viral). The method allows production of
XX mistletoe lectin, and its individual chains, in many different isoforms
XX and on a large scale, at any time of the year. Recombinant products are
XX free from toxins present in natural mistletoe extracts. This sequence
XX represents a consensus sequence of the mistletoe lectin described in the
XX specification.
SQ
SQ Sequence 533 AA;
SQ
SQ alignment_scores:
SQ Quality: 373.00 Length: 119
SQ Ratio: 3.885 Gaps: 1
SQ Percent Similarity: 80.672 Percent Identity: 65.546
SQ
SQ alignment_block:
SQ US-09-627-165B-13 x MAY25970 ..
SQ
SQ Align seg 1/1 to: MAY25970 from: 1 to: 533
SQ
SQ 1 GCCGATTCATTCGATCCTGTCGAGCGCTTCGCGCGCAATATPACAGTGG 50
SQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
SQ 168 AATGpnehsnroileutrlphg***Atgln***llelnserel 184
SQ
SQ 51 GCACTGCTTCACCAACATGTACATGCTGCAGCGTCGACAGCAATGGG 100
SQ ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
SQ 184 y***SerPheleuproasp***TyrMetleuclnleucclurhSerTTPG 201
SQ
SQ 101 GTGCACATATCCACCAAGTCACAGACTCCAGATGAGATTTTAAATCC 150

```

seq_name:	/SIDS1/scgdata/hold-geneseq/geneseqp-emb1/AAL1999.DAT.AAY25973
seq_documentation_block:	
ID	AAY25973 standard; protein; 533 AA.
XX	
AC	AAY25973;
XX	
DT	18-OCT-1999 (first entry)
XX	
DE	Mistletoe lectin protein consensus sequence 2.
XX	
KW	Mistletoe; lectin; antitumour; immunostimulant; A-chain; MIA; immunity;
KX	lysozyme 2B5 subunit; non-cytotoxic; T-cell activation; immune response
KV	lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW	cancer; cytotoxicity; antigen; isoform.
XX	
OS	Viscum album.
XX	
PID	Key Location/Qualifiers
FT	Misc-difference 15 /label= Asp, Glu
FT	Misc-difference 63 /label= Gly, Glu
FT	Misc-difference 66 /label= Ile, Val
FT	Misc-difference 75 /label= Leu, Ala
FT	Misc-difference 107 /label= Asp, Arg, none
FT	Misc-difference 113 /label= Asn, Thr
FT	Misc-difference 117 /label= Pro, Thr
FT	Misc-difference 134 /label= Asp, Glu
FT	Misc-difference 141 /label= Ser, Thr
FT	Misc-difference 145 /label= Phe, Tyr
FT	Misc-difference 152 /label= Thr, Ala
FT	Misc-difference 177 /label= Ala, Tyr
FT	Misc-difference 180 /label= Tyr, Asp
FT	Misc-difference 185 /label= Ala, Glu
FT	Misc-difference 191 /label= Val, Met
FT	Misc-difference 219

```

CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (1) activate T-cell and
CC lymphokine-producing macrophages, so stimulate immunity. (1) and its
CC fragments are used to treat uncontrolled cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains, in many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
CC represents a consensus sequence of the mistletoe lectin described in the
CC specification.
XX
XX
SQ Sequence 533 AA:

Alignment_scores:
    Quality: 373.00      Length: 119
    Ratio: 3.885        Gaps: 1
    Percent Similarity: 80.672      Percent Identity: 65.546

alignment_block:
US-09-627-165B-13 x AAY25973 ..

Align seg 1/1 to: AAY25973 from: 1 to: 533

1 GCCGATTCATCCCATGTCGTGGAGGCTCGCGGCAAAATTAAACATGG 50
||||| ||||| ||||| ||||| |||||
168 AlAaRphEhAsnProIleuTPrAArg**AArgIn**IleAAsnSerGI 184

51 GGAATCWTCTCCCAACCAATGTCATGCTACGTCAGCGGAGAGAGAGTTGG 100
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 y***SerPheLeuProAsp***TyrtleuGluIuengluInrserTrpG 201

101 GTGCAGATTCACACCAAGTCAGAGTCACAGAGATGGCATTTTAATACC 150
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 LygInGInSerThrGlnValGlnHisSerThrAspGlyValPheAsnAsn 217

131 CAATATGATTCAGATTTCCGCGCGGTAACTTTTGAGAGTACCAATGT 200
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
218 Pro***ArgleuAlaIle*****GlyAsnPheValThrLeu*****AsnVa 234

201 TCGCGAGCTGATCTCCACCTTGGCGATCATGTGTTTCGAATGACATGTC 250
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234 lArg**ValIleAlaSerleuAlaIleMetleuPheValGlySeliGluA 251

251 GGCAATCTCTCTCTCTCGACCAACCCTTCGCCGTGCTCTAAGTCTCGTC 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
251 rGProSerSerSerAspValArGTrYrTrpProleuValIleArgProVal 267

301 GTGATGTCGGGACACGATGTCACCTGGACACNTTTCGGAACACCGCGTGC 350
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
268 lIe.....AlaAspSpValThrCysSerAlaSerGluProThrValAr 282

351 CACGCTA 357
|||||
282 gIleVal 284

seq_name: /STDsI/gcgdata/hold-geneseq/geneseqp-embJ/AA1999.DAT:AA25976
seq_documentation_block:
ID AAY25976 standard; protein; 533 AA.
AC
AAAY25976:
XX
XX 18-OCT-1999 (first entry)
XX
XX Mistleoe lectin protein consensus sequence 3.
XX
XX Mistleoe lectin: antitumour; immunostimulant; A-chain; MLA; immunity;
XX ribosome 28S subunit; non-cytotoxic; T-cell activation; immune response;
XX lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX cancer; cytotoxicity; antigen; isoform.
KW

```

FT	Misc-difference	/label= Gly, Tyr
FT	500	
FT	Misc-difference	/label= Asn, Ser, Thr, Lys
FT	501	
FT	Misc-difference	/label= Ser, Gly
FT	502	
FT	Misc-difference	/label= Leu, Pro
FT	503	
FT	Misc-difference	/label= Ala, Met
FT	504	
FT	Misc-difference	/label= Met, Val
FT	533	
FT	Misc-difference	/label= Pro, Phe
XX		
XX	DE19804210-A1.	
XX		
PN	12-AUG-1999.	
XX		
PD		
XX	03-FEB-1998;	98DE-1004210.
PF		
XX	03-FEB-1998;	98DE-1004210.
PR		
XX	(BIOSYN ARZNEIMITTEL GMBH.	
XX		
PA		
PI	Morris P, Stiefel T, Voelter W, Welters P;	
XX		
XX	WPI; 1999-44535/38.	
DB		
XX	Preparation of mistletoe lectins in heterologous systems,	
PT	particularly for use as anticancer agents and immunostimulants	
PT		
XX		
DS	Claim 40; page 37-38; 78pp. German.	
XX		
XX	This invention describes a novel mistletoe lectin (I) and its fragments	
CC	which have antitumour and immunostimulatory activity. The A-chain (MA)	
CC	of the mistletoe lectin binds to, and inactivates, the 28S subunit of	
CC	ribosomes. Non-cytotoxic forms of (I) activate T-cell and	
CC	lymphokine-producing macrophages, so stimulate immunity. (I) and its	
CC	fragments are used to treat uncontrolled cell growth (particularly	
CC	cancers) and if they lack cytotoxicity, to increase the strength of the	
CC	immune response, particularly to a co-administered antigen	
CC	(tumour-associated, bacterial or viral). The method allows production of	
CC	mistletoe lectin, and its individual chains, in many different isoforms	
CC	and on a large scale, at any time of the year. Recombinant products are	
CC	free from toxins present in natural mistletoe extracts. This sequence	
CC	represents a consensus sequence of the mistletoe lectin described in the	
CC	specification.	
XX		
XX	Sequence	533 A1.
XQ		

```
151 CAAATTAAGATTGCAGATTTCGCCCGGTAACCTTGTGACGNTGACCAATGT 200
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
218 Pro**ArgLeuAlaIle*****GlyAsnPheValThrLeu***AsnVa 234
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 TCGCAGCGATCTCCAGCTTGCGATCATGTTGTTCGAATGCAAGTGTCTC 250
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234 IArg***ValIleAlaSerIleAlaIleMetLeuPheValCysGlyGluA 251
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
251 GGCCATTCTCTCTCTCGACACACCTTGGCGTCTCTCTTAAGGTCCGTC 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
251 rProSerSerSerAspValAlaArgTyrTrpProLeuValIleArgProVal 267
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 GTGATGATCGCGCCAGAGATGTCACACTNTTTCGAACCCACCGTGTG 350
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
268 ILe.....AlaAspAspValIThrCysSerAlaSerGluProThrValAr 282
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
351 CATCGTA 357
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
282 gIleVal 284

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:AAV25981

seq_documentation_block:
ID   AAV25981 standard; Protein; 256 AA.
XX
AC   AAV25981;
XX
DE   18-OCT-1999 (first entry)
XX
DE   Mistletoe lectin A2 protein fragment.
XX
XX
KW   Mistletoe lectin; antitumour; immunostimulant; A-chain; MIA; immunity;
KW   ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW   lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW   cancer; cytotoxicity; antigen; isoform; lectin A2.
XX
XX   Viscum album.
XX
XX   DE19804210-A1.
XX
XX   12-AUG-1999.
XX
XX   03-FEB-1998; 98DE-1004210.
XX
XX   03-FEB-1998; 98DE-1004210.
XX
XX   (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
XX   Morris P, Stiefel T, Voelter W, Welters P;
XX
XX   WPI: 1999-445335/38.
XX
XX   N-PSDB: AA209105.
XX
XX   Preparation of mistletoe lectins in heterologous systems,
XX   particularly for use as anticancer agents and immunostimulants
XX
XX   Claim 8; Fig 3B; 78pp; German.
XX
XX   This invention describes a novel mistletoe lectin (I) and its fragments
XX   which have antitumour and immunostimulatory activity. The A-chain (MIA)
XX   of the mistletoe lectin binds to, and inactivates, the 28S subunit of
XX   ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX   lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX   fragments are used to treat uncontrolled cell growth (particularly of the
XX   cancers) and if they lack cytotoxicity, to increase the strength of the
XX   immune response, particularly to a co-administered antigen
XX   (tumour-associated, bacterial or viral). The method allows production of
XX   mistletoe lectin, and its individual chains, in many different isoforms
XX   and on a large scale, at any time of the year. Recombinant products are
XX   free from toxins present in natural mistletoe extracts. This sequence
XX   represents a fragment of the mistletoe lectin A2 protein.
XX
SQ   Sequence 256 AA;
```

```
alignment_scores:
  Quality: 327.00      Length: 88
  Ratio: 4.247        Gaps: 0
  Percent Similarity: 87.500      Percent Identity: 73.864

alignment_block:
US-09-627-165b-13 x AAV25981  ..

Align seg 1/1 to: AAV25981 from: 1 to: 256

1  GCCAGATTCAATCCCATCTGAGAG3CTTGGCCCGGCAATTAACAGNG 50
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
169 AlaArgPheAsnProIleLeuTrpArgTyrArgGlnAspIleAsnSerGI 185
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 GAGGTCTTCGCCCAACATGTACATGCTGCAGCTGAGAGCATTTGGG 100
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
185 yGIuSerPheLeuProAspMetIleuGluLeuGluThrSerTrpG 202
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 GTCGACAAATCCACCCAAATGTCAGACATCCAAAGATGGCATTTTAATAC 150
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
202 IyGInGInSerThrGlnValGlnHisSerThrAspIValPheAsnVa 218
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151 CAAATTAAGATTGCAGATTTCGCCCGGTAACCTTGTGACGNTGACCAATGT 200
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 ProPheArgLeuAlaIleSerThrGlyAsnPheValThrLeuSerAsnVa 235
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 TCGCAGCGATCTCCAGCTTGCGATCATGTTGTTCGAATGCAAGTGTCTC 250
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
235 IArgSerValIleAlaSerIleAlaIleMetLeuPheValCysGlyGluA 252
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
251 GGCCATTCTCTCTCT 264
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
252 rProSerSerSer 256

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:AAV25984

seq_documentation_block:
ID   AAV25984 standard; Protein; 256 AA.
XX
AC   AAV25984;
XX
DE   18-OCT-1999 (first entry)
XX
DE   Mistletoe lectin A2 (variant) protein fragment.
XX
XX
KW   Mistletoe lectin; antitumour; immunostimulant; A-chain; MIA; immunity;
KW   ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW   lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW   cancer; cytotoxicity; antigen; isoform; lectin A2.
XX
XX   Viscum album.
XX
XX   DE19804210-A1.
XX
XX   12-AUG-1999.
XX
XX   03-FEB-1998; 98DE-1004210.
XX
XX   03-FEB-1998; 98DE-1004210.
XX
XX   (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
XX   Morris P, Stiefel T, Voelter W, Welters P;
XX
XX   WPI: 1999-445335/38.
XX
XX   N-PSDB: AA209108.
XX
XX   Preparation of mistletoe lectins in heterologous systems,
XX   particularly for use as anticancer agents and immunostimulants
XX
XX   Disclosure; Fig 6B; 78pp; German.
XX
PS
```

XX This invention describes a novel mistletoe lectin (I) and its fragments  
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)  
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of  
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and  
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its  
 CC fragments are used to treat uncontrolled cell growth (particularly  
 CC cancers) and if they lack cytotoxicity, to increase the strength of the  
 CC immune response, particularly to a co-administered antigen  
 CC (tumour-associated, bacterial or viral). The method allows production of  
 CC mistletoe lectin, and its individual chains, in many different isoforms  
 CC and on a large scale, at any time of the year. Recombinant products are  
 CC free from toxins present in natural mistletoe extracts. This sequence  
 CC represents a fragment of a mistletoe lectin A2 protein variant.

XX Sequence 256 AA:

alignment\_scores:                      Length:                      88  
                     Quality:                      327.00  
                     Ratio:                      4.247  
                     Gaps:                      0  
                     Percent Similarity:                      87.500  
                     Percent Identity:                      73.864

alignment\_block:

US-09-627-165B-13 x AAY25984 ..

Align seg 1/1 to: AAY25984 from: 1 to: 256

```

1  GCCAGATTCAATCCCATCTGTCGAGGCTCCGCGCAATTAACAGTGG 50
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
169  AlarphheasnProilleutrparglyrarginaspheasnsergi 185
51  GGACTCWTCTCCACCAATGATGATGCTGAGCGAGACGAGTGGG 100
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
185  ygluserpheleuprobaspmetlymethleuglndleugluthsertipg 202
101  GTCGCAATCCACCCAGATCCAGAGTCGCAAGATGCAATTTTAATACC 150
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
202  lynglnserthrnglnhisserthrpspalyvalpheasnasn 218
151  CAATTAAGATTGCAATTCGCGCGGTAACTTGTGACGNTGACCAATGT 200
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
219  ProheargleuhalleserthrnglyasnphelyalthrleuSerasnva 235
201  TCGGAGATGATCTCCAGCTTGCGCATGATGTGTGCAATGCGAGTGC 250
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
235  larservallilealsertleualilemetleuphevalcysglyglua 252
251  GGCAATCTCTCTCT 264
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
252  rtproserSerSer 256

```

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT: AAY25980

seq\_documentation\_block:

ID AAY25980 standard; Protein: 254 AA.

XX AAY25980:  
 AC  
 XX 18-OCT-1999 (first entry)  
 DE Mistletoe lectin A1 protein fragment.  
 XX  
 XX Mistletoe lectin; antitumour; immunostimulant; A-chain; MLA; immunity;  
 KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;  
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;  
 KW cancer; cytotoxicity; antigen; isoform; lectin A1.  
 XX  
 OS Viscum album.  
 XX  
 XX DE19804210-A1.  
 XX  
 PD 12-AUG-1999.

XX 03-FEB-1998; 98DE-1004210.  
 PF  
 XX 03-FEB-1998; 98DE-1004210.  
 PR  
 XX (BIO5-) BIOSYN ARZNEIMITTEL GMBH.  
 PA  
 PI Morris P, Stiefel Y, Voelter W, Welters P;  
 XX  
 DR WPI: 1999-445335/38.  
 DR N-PSDB: AA209104.

PT Preparation of mistletoe lectins in heterologous systems,  
 particularly for use as anticancer agents and immunostimulants  
 PR  
 XX  
 XX  
 PS Disclosure: Fig 2B: 78pp: German.

XX This invention describes a novel mistletoe lectin (I) and its fragments  
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)  
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of  
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and  
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its  
 CC fragments are used to treat uncontrolled cell growth (particularly  
 CC cancers) and if they lack cytotoxicity, to increase the strength of the  
 CC immune response, particularly to a co-administered antigen  
 CC (tumour-associated, bacterial or viral). The method allows production of  
 CC mistletoe lectin, and its individual chains, in many different isoforms  
 CC and on a large scale, at any time of the year. Recombinant products are  
 CC free from toxins present in natural mistletoe extracts. This sequence  
 CC represents a fragment of the mistletoe lectin A1 protein.

XX Sequence 254 AA:

alignment\_scores:                      Length:                      88  
                     Quality:                      317.00  
                     Ratio:                      4.342  
                     Gaps:                      0  
                     Percent Similarity:                      82.955  
                     Percent Identity:                      71.591

alignment\_block:

US-09-627-165B-13 x AAY25980 ..

Align seg 1/1 to: AAY25980 from: 1 to: 254

```

1  GCCAGATTCAATCCCATCTGTCGAGGCTCCGCGCAATTAACAGTGG 50
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
167  AlarphheasnProilleutrparglyrarginleuhsnsergi 183
51  GGACTCWTCTCCACCAATGATGATGCTGAGCGAGACGAGTGGG 100
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
183  yalaserpheleuprobaspvallymethleuglndleugluthsertipg 200
101  GTCGCAATCCACCCAGATCCAGAGTCGCAAGATGCAATTTTAATACC 150
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
200  lynglnserthrnglnhisserthrpspalyvalpheasnasn 216
151  CAATTAAGATTGCAATTCGCGCGGTAACTTGTGCAATGCGAGTGC 200
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
217  Proleargleuhalleproproliysnphelyalthrleuhrsasnva 233
201  TCGGAGATGATCTCCAGCTTGCGCATGATGTGTGCAATGCGAGTGC 250
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
233  larspvalillealsertleualilemetleuphevalcysglyglua 250
251  GGCAATCTCTCTCT 264
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
250  rtproserSerSer 254

```

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT: AAY25983

seq\_documentation\_block:

ID AAY25983 standard; Protein: 254 AA.

XX



## alignment\_scores:

Quality: 311.00 Length: 85  
Ratio: 4.380 Gaps: 0  
Percent Similarity: 83.529 Percent Identity: 71.765

## alignment\_block:

US-09-627-165b-13 x AAW64659 ..

Align seg 1/1 to: AAW64659 from: 1 to: 252

```

1 GCCAGATTCAATCCCATCMTGTGGAGCGCTCCGCCGCAAAATTAACAGTCG 50
|||||
168 AlaArgPheAsnProIleuTrpArgAlaArgGlnTyrIleAsnSerG1 184
51 GGAGTCNTCTCCACCAACATGTACATGCTCGAGCGTGGAGAGAGATGGG 100
| ||| |||:|||||
184 yAlaSerPheLeuProAspValIyrMetLeuGlnLeuGlnTyrSerTrpG 201
101 GTCCGACATCCACCCAAATGTCAGACAGTCCAGAGATGATTTTAATACC 150
||:|||||
201 IyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsnAsn 217
151 CAATTAAGATTGACATTTCCCGCGTAACTTTTGACAGTACAGCAATGT 200
|||||
218 ProIleArgLeuAlaIleProProGlyAsnPheValThrLeuThrAsnVa 234
201 TCGCGACGTGATCTCCAGCTTGCGCATCATGTTGTCGAATGACATGTC 250
|||||
234 IArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGluA 251
251 GGCCA 255
|||||
251 rGPro 252

```

seq\_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:AAW64661

seq\_documentation\_block:

ID AAW64661 standard; Protein; 252 AA.

```

XX AC AAW64661;
XX DT 23-OCT-1998 (first entry)
XX DE Mistletoe rMLA variant protein.
XX KW Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
KW intracellular; processing module; protease recognition; targeting module;
KW internalisation; treatment; disorder; cell proliferation; activation;
KW autoimmune disease; allergy; tumour; ricin; translocation; ss.
XX OS Viscum album.
XX FT key Location/Qualifiers
FT Protein 1..252
FT /note= "partial"
XX PN MO9829540-A2.
XX PD 09-JUL-1998.
XX PF 02-JAN-1998; 98WO-EP00009.
XX PR 02-JAN-1997; 97EP-0100012.
XX PA (BRAL-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.
XX PI Eck J, Schmidt A, Zinke H;
XX DR WPI: 1998-388122/33.
XX N-PSDB: AAV51343.
XX

```

PF Nucleic acid encoding fusion protein containing mistletoe lectin A  
PT chain - useful for treatment of proliferative and autoimmune  
PS diseases, allergies and tumours  
XX Disclosure; Fig 11a'; 115pp; German.

CC This sequence encodes a variant mistletoe lectin A-chain, rMLA. This  
CC sequence can be used in the construction of a fusion protein which  
CC comprises an effector module that is cytotoxic intracellularly, a  
CC processing module covalently bonded to the effector module and  
CC containing a protease recognition sequence, and a targeting module  
CC covalently bonded to the processing module, able to bind specifically to  
CC the surface of a cell so as to mediate internalisation of the fusion  
CC protein. Such a fusion protein can be used for treating disorders  
CC involving proliferation and/or elevated activation of cells, especially  
CC autoimmune disease, allergy and tumours. The proteins can be administered  
CC e.g. by injection or topically but especially by intravenous injection,  
CC at 1 ng to 500 mu g/kg/day, or for ex vivo use at 1 pg to 500 ng/mL.  
CC Fusion proteins can develop toxic activity in a wide range of target  
CC cells. The processing module prevents extracellular dissociation, and  
CC fusion proteins based on mistletoe lectin A-chain are far more active  
CC than those based on ricin and do have the associated problems of  
CC non-specific toxicity. The protein may be expressed in a non-glycosylated  
CC form that does not bind to sugar receptors in the liver, and which has a  
CC long half-life in the blood, where the mistletoe lectin B-chain is used,  
CC it actively assists in translocation of the ML A-chain from the  
CC endoplasmic reticulum to the cytoplasm.

Sequence 252 AA:

alignment\_scores:  
Quality: 311.00 Length: 85  
Ratio: 4.380 Gaps: 0  
Percent Similarity: 83.529 Percent Identity: 71.765

## alignment\_block:

US-09-627-165b-13 x AAW64661 ..

Align seg 1/1 to: AAW64661 from: 1 to: 252

```

1 GCCAGATTCAATCCCATCMTGTGGAGCGCTCCGCCGCAAAATTAACAGTCG 50
|||||
167 AlaArgPheAsnProIleuTrpArgAlaArgGlnTyrIleAsnSerG1 183
51 GGAGTCNTCTCCACCAACATGTACATGCTCGAGCGTGGAGAGAGATGGG 100
| ||| |||:|||||
183 yAlaSerPheLeuProAspValIyrMetLeuGlnLeuGlnTyrSerTrpG 200
101 GTCCGACATCCACCCAAATGTCAGACAGTCCAGAGATGATTTTAATACC 150
||:|||||
200 IyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsnAsn 216
151 CAATTAAGATTGACATTTCCCGCGTAACTTTTGACAGTACAGCAATGT 200
|||||
217 ProIleArgLeuAlaIleProProGlyAsnPheValThrLeuThrAsnVa 233
201 TCGCGACGTGATCTCCAGCTTGCGCATCATGTTGTCGAATGACATGTC 250
|||||
233 IArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGluA 250
251 GGCCA 255
|||||
251 rGPro 251

```





314 gileval 316

seq\_name: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:US-08-776-059-39

seq\_documentation\_block:

Sequence 39, Application US/08776059B  
Patent No. 6271368  
GENERAL INFORMATION:  
APPLICANT: LENTZEN, Hans  
APPLICANT: ECK, Jürgen  
APPLICANT: BAUR, Axel  
APPLICANT: ZINKE, Holger  
TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)  
FILE REFERENCE: 674503-2003  
CURRENT FILING DATE: 1999-06-19  
EARLIER FILING DATE: 1996-06-25  
EARLIER FILING DATE: 1996-06-25  
EARLIER APPLICATION NUMBER: PCT/EP96/02273  
EARLIER FILING DATE: 1995-06-26  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 39  
LENGTH: 235  
TYPE: PRT  
ORGANISM: Viscum album  
US-08-776-059-39

alignment\_scores:

Quality: 311.00 Length: 85  
Ratio: 4.380 Gaps: 0  
Percent Similarity: 83.529 Percent Identity: 71.765

alignment\_block:

US-09-627-165b-13 x US-08-776-059-39 ..

Align seg 1/1 to: US-08-776-059-39 from: 1 to: 235

```
1 GCCAGATTCAATCCCATGTCGAGAGCTTCGCCGCGCAATTAACAGTGG 50
|||||
150 AIAARGPHEASNPRIILEUTRPGALARGLINTYRILEASNSERTGL 166
51 GGAGTCNTCTCCACCAACATGACATGCTCGAGCTCGAGACGAGTGGG 100
|||||
166 YALASERPHLEUPROASPVALLYRMETLEUGLDELGLUTHSERTPG 183
101 GTCCAGATCCACCCAGCTCGAGCACTCCAGAGATGCATTTTAATACC 150
|||||
183 IYGLINLSERTHGINVALGINHLSERTHRASPGLYVALPHEASNSN 199
151 CAATTAAGATTGCGATTTCGCGCGSTAACTTTGTGAGAGTGCATATG 200
|||||
200 PROILEARGLEUALALEPROPIGGLYASNPHEVALTHIRLEUTHRASNA 216
201 TCCGAGATGATCTCCAGCTTGCGGATCATGTTTGCATGCGAGTGGT 250
|||||
216 IARGASPVALLIETALASERLEUALILEMETLEUPHERVALCYSGLYLUA 233
251 GGCCA 255
|||||
233 rgrPro 234
```

seq\_name: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:US-08-776-059-31

seq\_documentation\_block:

Sequence 31, Application US/08776059B  
Patent No. 6271368  
GENERAL INFORMATION:  
APPLICANT: LENTZEN, Hans  
APPLICANT: ECK, Jürgen  
APPLICANT: BAUR, Axel  
APPLICANT: ZINKE, Holger

TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)

FILE REFERENCE: 674503-2003  
CURRENT APPLICATION NUMBER: US/08/776, 059B  
CURRENT FILING DATE: 1999-06-19  
EARLIER APPLICATION NUMBER: PCT/EP96/02273  
EARLIER FILING DATE: 1996-06-25  
EARLIER APPLICATION NUMBER: 95109949.8  
EARLIER FILING DATE: 1995-06-26  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 31  
LENGTH: 253  
TYPE: PRT  
ORGANISM: Viscum album  
US-08-776-059-31

alignment\_scores:

Quality: 311.00 Length: 85  
Ratio: 4.380 Gaps: 0  
Percent Similarity: 83.529 Percent Identity: 71.765

alignment\_block:

US-09-627-165b-13 x US-08-776-059-31 ..

Align seg 1/1 to: US-08-776-059-31 from: 1 to: 253

```
1 GCCAGATTCAATCCCATGTCGAGAGCTTCGCCGCGCAATTAACAGTGG 50
|||||
168 AIAARGPHEASNPRIILEUTRPGALARGLINTYRILEASNSERTGL 184
51 GGAGTCNTCTCCACCAACATGACATGCTCGAGCTCGAGACGAGTGGG 100
|||||
184 YALASERPHLEUPROASPVALLYRMETLEUGLDELGLUTHSERTPG 201
101 GTCCAGATCCACCCAGCTCGAGCACTCCAGAGATGCATTTTAATACC 150
|||||
201 IYGLINLSERTHGINVALGINHLSERTHRASPGLYVALPHEASNSN 217
151 CAATTAAGATTGCGATTTCGCGCGSTAACTTTGTGAGAGTGCATATG 200
|||||
218 PROILEARGLEUALALEPROPIGGLYASNPHEVALTHIRLEUTHRASNA 234
201 TCCGAGATGATCTCCAGCTTGCGGATCATGTTTGCATGCGAGTGGT 250
|||||
234 IARGASPVALLIETALASERLEUALILEMETLEUPHERVALCYSGLYLUA 251
251 GGCCA 255
|||||
251 rgrPro 252
```

seq\_name: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:US-08-378-761A-77

seq\_documentation\_block:

Sequence 77, Application US/08378761A  
Patent No. 5635384  
GENERAL INFORMATION:  
APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D  
APPLICANT: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk



```
261 roProSerSerGlnPheSer.....LeuLeuLeuArgProVal 273
301 GTGATGCGGGCCAGCATGTCACCGCATTTTCGGAACCCACCGTGG 350
274 ValProAsnPhenAsnAlaAspValCysMetAspProGluProIleValAr 290
351 CATCGTA 357
290 gileVal 292

seq_name: /cgn2.6/prodata/2/1aa/5b_COMB.pep:us-08-356-786-10

seq_documentation_block:
: Sequence 10, Application US/08356786
: Patent No. 5877305
: GENERAL INFORMATION:
: APPLICANT: Huston, James S.
: APPLICANT: Oppermann, Hermann
: APPLICANT: Houston, L. L.
: APPLICANT: Ring, David B.
: TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
: TITLE OF INVENTION: Marker
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
: STREET: Exchange Place, 53 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/356,786
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/831,967
: FILING DATE: 06 FEB 1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Pitcher, Edmund R.
: REGISTRATION NUMBER: 27,829
: REFERENCE/DOCKET NUMBER: CRP-053
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 248-7000
: TELEFAX: (617) 248-7100
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 534 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-356-786-10

alignment_scores:
: Quality: 159.00 Length: 119
: Ratio: 2.092 Gaps: 2
: Percent Similarity: 63.866 Percent Identity: 33.613

alignment_block:
US-09-627-165b-13 x US-08-356-786-10 ..
Align seq 1/1 to: US-08-356-786-10 from: 1 to: 534
1 GCGAGATTCATCCCATGTCGAGGCTTCGCGCAATTAAGAGTGG 50
182 AlaArgPheGlnIleArgIleGluGluMetLeuArgIleArgIleArgTyrAs 198
51 GGAAGTCATCTCCACCAAGCATGTACATGCTCGAGCTGGAGAGAGATTGGG 100
```

```
198 nArgArgSerAlaProAsnProSerValIleIleIleGluIleAsnSerTrrp 215
101 GTGCAACATTCACCAAGTCACCGACAGTCACAGATGATGATTTTAATACC 150
215 LysArgLeuSerThrAlaIleGlnGlnIleSerAsnGlnGlnIleAlaPheAlaSer 231
151 CAAATAGATTGACAGATTTCGCCGCTTACCTTGTGACGNTGACCAATGT 200
232 ProIleGlnLeuGlnArgArgAsnGlySerIlySpheSerValTyrAspVal 248
201 TCGCACGTCATCTCCACGCTTGGCGATCATGTGTGTCGAATGACAGTGC 250
248 lSerIleLeuIleProIleIleAlaIleLeuMetValTyrArgCysAlaPro 265
251 GGCATTCCTCTCTGCGACACACCTTGCCTGCGCGCTGCTTCCGAAACCGTGG 300
265 roProSerSerGlnPheSer.....LeuLeuIleArgProVal 277
301 GTGATGCGGGCCAGCATGTCACCGCATTTTCGGAACCCACCGTGG 350
278 ValProAsnPhenAsnAlaAspValCys.....MetAspProGluIleG 292
351 CATCGTA 357
292 nleuVal 294

seq_name: /cgn2.6/prodata/2/1aa/5a_COMB.pep:us-07-901-707-1

seq_documentation_block:
: Sequence 1, Application US/07901707
: Patent No. 5376546
: GENERAL INFORMATION:
: APPLICANT: Bernhard, Susan L.
: APPLICANT: Better, Marc D.
: APPLICANT: Carroll, Steve F.
: APPLICANT: Lane, Julie A.
: TITLE OF INVENTION: Materials Comprising and Methods of
: TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins
: NUMBER OF SEQUENCES: 57
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
: STREET: Two First National Plaza, 20 South Clark
: STREET: Street
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60603
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/901,707
: FILING DATE: 19920619
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/787,567
: FILING DATE: 04-NOV-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5376546and, Greta E.
: REGISTRATION NUMBER: 35,302
: REFERENCE/DOCKET NUMBER: 27129/30910
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 346-5750
: TELEFAX: (312) 984-5750
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 267 amino acids
: TYPE: AMINO ACID
```

```

;
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   US-07-901-707-1

alignment_scores:
    Quality: 138.00      Length: 87
    Ratio: 2.379         Gaps: 0
    Percent Similarity: 66.667      Percent Identity: 35.632

alignment_block:
US-09-627-165b-13 x US-07-901-707-1 ..

Align seg 1/1 to: US-07-901-707-1 from: 1 to: 267

1   GCCAGATTCATCCATCCTGAGGCTTCCGCGCAATTAACAGTGG 50
|||||
179  AlAaRgPheGlnTyrIleGlnGlyLueMetArGThrArgIleArGTYrAs 195
51  GGAGTCNTCTCCACCAACATGTACATGTGCGAGCGAGAGTGGG 100
|||||
195  nArGArGSeRAlaPrOaSPProSeRValIleThrLeuGlnsnsertPg 212
101  GTGACATTCACCCAGTCACAGCTCCAGAGATGGCATTTTAAATAC 150
|||||
212  lYArGLeuSeRThrAlaIleGlnGluSeRAsnGlnGlyAlaPheAlaSeR 228
151  CAATTAAGATTGCAGATTTCGCCGCTTAACCTTTGTGACGNTGACCAATCT 200
|||||
229  ProIleGlnLeuGlnArGArGAsnGlySeRlYsPheSeRValTYrAsPvA 245
201  TCGGACGTGATCTCCAGCTTGCGGATCATGTTGTCGATGACAGTGTC 250
|
245  lSerIleuLeuProIleAlaIleGlnGluSeRValTYrArGcYsAlaPrOP 262
251  GGCCATTCTCC 261
|||
262  rOProSeRSeR 265

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-07-988-430-1

seq_documentation_block:
; Sequence 1, Application US/07988430
;
;   GENERAL INFORMATION:
;   APPLICANT: Bernhard, Susan L.
;   APPLICANT: Better, Marc D.
;   APPLICANT: Carroll, Stephen F.
;   APPLICANT: Lane, Julie A.
;   APPLICANT: Lei, Shau-Ping
;   TITLE OF INVENTION: Materials Comprising and Methods of
;   TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins
;   NUMBER OF SEQUENCES: 101
;   CORRESPONDENCE ADDRESSES:
;   ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
;   STREET: Two First National Plaza, 20 South Clark
;   STREET: Street
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: USA
;   ZIP: 60603
;
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/988,430
;   FILING DATE: 19921209
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/901,707
```

```

;
;   FILING DATE: 19-JUN-1992
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/787,567
;   FILING DATE: 04-NOV-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: No. 541620zand, Greta E.
;   REGISTRATION NUMBER: 35302
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (312) 346-5750
;   TELEFAX: (312) 984-9740
;   TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 267 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   US-07-988-430-1

alignment_scores:
    Quality: 138.00      Length: 87
    Ratio: 2.379         Gaps: 0
    Percent Similarity: 66.667      Percent Identity: 35.632

alignment_block:
US-09-627-165b-13 x US-07-988-430-1 ..

Align seg 1/1 to: US-07-988-430-1 from: 1 to: 267

1   GCCAGATTCATCCATCCTGAGGCTTCCGCGCAATTAACAGTGG 50
|||||
179  AlAaRgPheGlnTyrIleGlnGlyLueMetArGThrArgIleArGTYrAs 195
51  GGAGTCNTCTCCACCAACATGTACATGTGCGAGCTGAGACAGTGGG 100
|||||
195  nArGArGSeRAlaPrOaSPProSeRValIleThrLeuGlnsnsertPg 212
101  GTGACATTCACCCAGTCACAGCTCCAGAGATGGCATTTTAAATAC 150
|||||
212  lYArGLeuSeRThrAlaIleGlnGluSeRAsnGlnGlyAlaPheAlaSeR 228
151  CAATTAAGATTGCAGATTTCGCCGCTTAACCTTTGTGACGNTGACCAATGT 200
|||||
229  ProIleGlnLeuGlnArGArGAsnGlySeRlYsPheSeRValTYrAsPvA 245
201  TCGGACGTGATCTCCAGCTTGCGGATCATGTTGTCGATGACAGTGTC 250
|
245  lSerIleuLeuProIleAlaIleGlnGluSeRValTYrArGcYsAlaPrOP 262
251  GGCCATTCTCC 261
|||
262  rOProSeRSeR 265

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-218-303-16

seq_documentation_block:
; Sequence 16, Application US/08218303
;
;   GENERAL INFORMATION:
;   APPLICANT: Kara, Bhupendra V.
;   APPLICANT: Hockney, Robert C.
;   APPLICANT: Filton, John B.
;   TITLE OF INVENTION: FERMENTATION PROCESS
;   NUMBER OF SEQUENCES: 23
;   CORRESPONDENCE ADDRESSES:
;   ADDRESSEE: Cushman, Darby & Cushman
;   STREET: 1615 L Street, N.W.
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20036-5601
```



251 GGCATCTCC 261  
||| |||  
262 rProSerSer 265

seq\_name: /cgn2\_6/ptodata/2/1aa/5A-COMB.pep:US-08-488-113B-1

seq\_documentation\_block:

; Sequence 1, Application US/08488113B

; Patent No. 5744580

; GENERAL INFORMATION:

; APPLICANT: Better, Marc D.

; APPLICANT: Carroll, Stephen F.

; APPLICANT: Studnika, Gary M.

; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

; TITLE OF INVENTION: Proteins

; NUMBER OF SEQUENCES: 169

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McAndrews, Held & Malloy, Ltd.

; STREET: 500 West Madison Street, 34th floor

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60661

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/488,113B

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/425,336

; FILING DATE: 18-APR-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/064,691

; FILING DATE: 12-MAY-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/988,430

; FILING DATE: 09-DEC-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/901,707

; FILING DATE: 19-JUN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/787,567

; FILING DATE: 04-NOV-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: McNicholas, Janet M.

; REGISTRATION NUMBER: 32,918

; REFERENCE/DOCKET NUMBER: 110220S07/200-70.P3.C2A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312/707-8889

; TELEFAX: 312/707-9155

; TELEX: 650 388-1248

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 267 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-488-113B-1

alignment\_scores:

Quality: 138.00 Length: 87

Ratio: 2.379 Gaps: 0

Percent Similarity: 66.667 Percent Identity: 35.632

alignment\_block:

US-09-627-165b-13 x US-08-488-113B-1 ..

Align seg 1/1 to: US-08-488-113B-1 from: 1 to: 267

1 GCCAGATTCATCCCATCCTGAGAGCTTCCGCGCAATTACAGTGC 50  
|||||  
179 AAlaArgPheGlnTyrIleGluGlyGluMetArgThrArgIleArgTyrAs 195  
::: |||  
51 GGAGTCNTCTCCACCAACATGTACATGCTCGAGCTGAGAGAGAGTGGG 100  
::: |||  
195 nArgTgSerAlaProAspProSerValIleThrIleGlnAsnSerTyrPc 212  
101 GTCCGACATCCACCAAGTCCAGAGTCCAGAGATGGCTTTTATATAC 150  
|||||  
212 lArgLeuSerThrAlaIleGlnIleSerAsnGlnGlyAlaIleAlaSer 228  
151 CAATTAAGATTGCAGATTTCGCCGCTTAACCTTGTGACGATGACCAATGT 200  
|||||  
229 ProIleGlnLeuGlnArgArgAsnGlySerLysPheSerValTyrAspVa 245  
201 TCGCAGCGATCTCCAGCTTGCGATCATGTTGTTGCAAGACAGTGTGC 250  
|  
245 lSerIleuIleProlIleIleAlaLeuMetValTyrArgCysAlaProP 262  
251 GGCATCTCC 261  
||| |||  
262 rProSerSer 265

seq\_name: /cgn2\_6/ptodata/2/1aa/5A-COMB.pep:US-08-477-484B-1

seq\_documentation\_block:

; Sequence 1, Application US/08477484B

; Patent No. 5756699

; GENERAL INFORMATION:

; APPLICANT: Better, Marc D.

; APPLICANT: Carroll, Stephen F.

; APPLICANT: Studnika, Gary M.

; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

; TITLE OF INVENTION: Proteins

; NUMBER OF SEQUENCES: 169

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McAndrews, Held & Malloy, Ltd.

; STREET: 500 West Madison Street, 34th floor

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60661

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,484B

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/425,336

; FILING DATE: 18-APR-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/064,691

; FILING DATE: 12-MAY-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/988,430

; FILING DATE: 09-DEC-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/901,707

; FILING DATE: 19-JUN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/787,567

; FILING DATE: 04-NOV-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: McNicholas, Janet M.

; REGISTRATION NUMBER: 32,918

; REFERENCE/DOCKET NUMBER: 110220S07/200-70.P3.C2A

; TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-477-484B-1

Alignment\_scores:  
Quality: 138.00 Length: 87  
Ratio: 2.379 Gaps: 0  
Percent Similarity: 66.667 Percent Identity: 35.632

alignment\_block:  
US-09-627-165B-13 x US-08-477-484B-1 ..

Align seg 1/1 to: US-08-477-484B-1 from: 1 to: 267

```
1 GCCAGATTCAATCCATGTCGTGGAGCTTCGCCGCAATTAACAGTGG 50
|||||::: ||| ::::||| ::::|||:::
179 Alarphgelnlyrlllelgluglylmetarphrghlearglyras 195
51 GGAGTCCTCTCCACCAACATGATGATGCTGAGCTGGAGACAGATTGG 100
::: ||| ||::: :::: ||| ||::: ||| |||
195 nararqseralaproasproserVallethrleugluanserrtgs 212
101 GTGCACATCCACCAAGTCACAGTCACAGATGGCATTTTAAATACC 150
||||| ||||| ::::|||::: ||| |||
212 lyargleuserthrAlallelglngluserAanglnglyAlaphealaser 228
151 CAATAGATTGCAGATTCCCGCGTAACCTTGTGACGNTGAGCAATGT 200
||||| ||||| ||::: |||::: |||:::
229 ProileglnleuglnarqarqasnglyserlysheserValtyraspya 245
201 TCCGAGCTGATCTCCAGCTGGCGATGATGTTGTGATGCAAGTGGTC 250
|::: |||::: |||::: |||::: |||::: |||:::
245 lserlleuilelproillelleleleuMetValtyrarqysalapro 262
251 GGCCATTCTCC 261
||| |||
262 roProSerSer 265
```

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-646-360-1

seq\_documentation\_block:

; Sequence 1, Application US/08646360  
; Patent No. 5837491  
; GENERAL INFORMATION:  
; APPLICANT: Belter, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Staudika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 173  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/646,360  
; FILING DATE: 13-MAY-1996

CLASSIFICATION: 530  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/05348  
; FILING DATE: 12-MAY-1994  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 200-70,P4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-646-360-1

alignment\_scores:  
Quality: 138.00 Length: 87  
Ratio: 2.379 Gaps: 0  
Percent Similarity: 66.667 Percent Identity: 35.632

alignment\_block:  
US-09-627-165B-13 x US-08-646-360-1 ..

Align seg 1/1 to: US-08-646-360-1 from: 1 to: 267

```
1 GCCAGATTCAATCCATGTCGTGGAGCTTCGCCGCAATTAACAGTGG 50
|||||::: ||| ::::||| ::::|||:::
179 Alarphgelnlyrlllelgluglylmetarphrghlearglyras 195
51 GGAGTCCTCTCCACCAACATGATGATGCTGAGCTGGAGACAGATTGG 100
::: ||| ||::: :::: ||| ||::: ||| |||
195 nararqseralaproasproserVallethrleugluanserrtgs 212
101 GTGCACATCCACCAAGTCACAGTCACAGATGGCATTTTAAATACC 150
||||| ||||| ::::|||::: ||| |||
212 lyargleuserthrAlallelglngluserAanglnglyAlaphealaser 228
151 CAATAGATTGCAGATTTCGCCGTAACCTTGTGACGNTGAGCAATGT 200
||||| ||||| ||::: |||::: |||:::
229 ProileglnleuglnarqarqasnglyserlysheserValtyraspya 245
201 TCCGAGCTGATCTCCAGCTGGCGATGATGTTGTGATGCAAGTGGTC 250
|::: |||::: |||::: |||::: |||::: |||:::
245 lserlleuilelproillelleleleuMetValtyrarqysalapro 262
251 GGCCATTCTCC 261
||| |||
262 roProSerSer 265
```

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-338-793D-61

seq\_documentation\_block:

; Sequence 61, Application US/08338793D  
; Patent No. 5840521  
; GENERAL INFORMATION:



APPLICANT: Barth, Peter Thomas  
TITLE OF INVENTION: VECTOR  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN DABRY CUSHMAN  
ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF  
ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
COMPUTER: IBM PC/XT/AT Compatibles  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Microsoft Word or ASCII editors  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/338, 793D  
FILING DATE: 08-No. 5840521-94  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842, 081  
FILING DATE: 26-Feb-92  
CLASSIFICATION: 435  
APPLICATION NUMBER: 9104017.0  
FILING DATE: 26-Feb-91  
APPLICATION NUMBER: 9109188.4  
FILING DATE: 29-Apr-91  
ATTORNEY/AGENT INFORMATION:  
NAME: Kokulis, Paul N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: DDB/9901/215431/TCM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
US-08-338-793D-61

alignment\_scores:  
Quality: 138.00 Length: 87  
Ratio: 2.379 Gaps: 0  
Percent Similarity: 66.667 Percent Identity: 35.632

Alignment\_block:  
US-09-627-165b-13 x US-08-338-793D-61 ..

Align seg 1/1 to: US-08-338-793D-61 from: 1 to: 267

```
1 GCGAGATTCAATCCATCCTGTCGAGGCTCGCGCAATTAACAGTCG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179 AlaaTgphgIntTtYlleglUglUmeLarGthraTgIleatGfYras 195
51 GGAGTCTCTCCACCAAAACATATGATGCTGAGCTGGAGACGAGTGGG 100
::: ||| |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
195 nargTgSerAlaProaSpProSeTValIleThLeuLInuSnsSerTpg 212
101 GTGGAGATTCACCCCAAGTCACGAGTCACAGATGCGATTGTAATACC 150
||||| |||||||:|||||:|||||:|||||:|||||:|||||:|||||
212 lYarTgLeuSerThraIaIleGInglUSeTAsngInglYAlaPheHlaSer 228
151 CAATAAGATTGAGATTTCGCCCGCTAACTTTGGAGCAGTACGAATCT 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
229 ProIlleGInleuGInaTgArGasnGlySerLysPheSerValTyrAsp 245
201 TCGGAGCGTATCTCCAGCTTGCGATCATGTTGTTGCAATCAGTGTCT 250
```

245 lSerIleLeuIleProIleIleAlaLeuMeTValTyrArGysAlaIaPro 262  
251 GCCCATCTCTCC 261  
262 roProSerSer 265

seq\_name: /cgn2\_6/plodata/2/1aa/5B\_COMB.pep:us-08-839-765-1

seq\_documentation\_block:  
Sequence 1, Application US/08839765  
Patent No. 6146631  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Stuhnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,765  
FILING DATE: 15-APR-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,913  
REFERENCE/DOCKET NUMBER: 110220509/200-70.P3.C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-839-765-1

alignment\_scores:  
Quality: 138.00 Length: 87  
Ratio: 2.379 Gaps: 0  
Percent Similarity: 66.667 Percent Identity: 35.632

Alignment\_block:

US-09-627-165B-13 x US-08-839-765-1 ..

Align seq 1/1 to: US-08-839-765-1 from: 1 to: 267

```

1 GCCGATTTCATCCCATCNCNTGTGAGGCTTCGCCGCAATTAACAGTGG 50
  |||||::: ||| ::::| | |::|::: |
179 AlaArgPheGlnTyrIleGlnGluMetArgThrArgIleArgTyrAs 195
  :::: ||| ||::: ::::: |||||::: |||:::
51 GGAAGTCNTCTCCACCAACATGTACATGCTCGAGCTGAGAGAGTGGG 100
  :::: ||| ||::: ::::: |||||::: |||||
195 nArgArgSerAlaProAspProSerValIleThrLeuGluAsnSerTrpG 212
  :::: |||||::: |||||::: |||||::: |||||
101 GTCGACATGCAACCAAGTCCAGCAGTCCAGAGATGGCATTTTAATACC 150
  ||||| |||||::: |||||::: |||||::: |||:::
212 LysArgLeuSerThrAlaIleGlnGluSerAsnGlnGlyAlaPheAlaSer 228
  151 CAAATAAGATTGCAGATTTCGCCGCTAACTTTGTGACGNTGAGCAATCT 200
  ||::: ||||| ||::: |||||::: |||||::: |||
229 ProIleGlnLeuGlnArgArgAsnGlySerLysPheSerValTyrAspVa 245
  201 TGGCGACGATGATCCAGCTGGCGATCATGTTGTTGAATGCAAGTGGTC 250
  |::: ||::: ||::: ||::: ||::: ||::: ||:::
245 LserIleLeuIleProIleIleAlaLeuMetValTyrArgCysAlaPro 262
  251 GGCCATTTCCTCC 261
  ||| |||
262 roProSerSer 265
```

us-09-627-165b-13.rpt

OM of: US-09-627-165B-13 to: PIR\_71: \* out\_format : p1s  
Date: Jul 30, 2002 3:55 PM

About: Results were produced by the genCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:  
-model=frame+ n2p.model

```

MODEL=frame+2p.model -DEV=zip
-O/-c/gn2.1/USPPO.apool/US09627165/runa1.30072002.151139.3335/app.query.fasta_1.1.102
-DB=PI_1 -DEMT=fastin -SUFFIX=rrp -GAPOP=12.000 -GAPEXT=4.000
-MIMMACH=0.100 -LOOPEL=0.000 -LOOPEXT=0.500 -GAPOP=6.000
-OGAPEXT=0.500 -XGAPOP=10.000 -YGAPOP=6.000
-OGAPEXT=0.500 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-FCGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -TRANS=humaa40.cdl
-DELETEXT=7.000 -START=1 -MATRIX=blsnum62 -TRANS=humaa40.cdl
-list=4-5 -DCCALIGN=200 -THR.SCORE=pccl -THR.MAX=10 -THR.MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pls -NCM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09627165.ecgn1.112
-NCPU=6 -ICPU=3 -LONGJOB -DEV.TIMEOUT=120 -WARN.TIMEOUT=30
-NO.XLIFY -WAIT -THREADS=1

```

Search Information block:  
Query: US-09-627-165B-13

```
Database: PIR/1.1
Database sequences: 283138
Database length: 96089334
Search time (sec): 72.560000
```

score\_11st:

Sequence	Strd	Orig	ZScore	EScore	Len	Docname
P1R2:PD0018	+	317.00	619.58	8.0e-27	234	musclelec lectin I A chain - V
P1R2:RLCSD	+	178.00	332.62	2.6e-11	576	richin D precursor - castor bean
P1R1:RLCSAG	+	177.00	334.81	3.4e-11	564	agglutinin precursor - castor b
P1R1:TRLSA	+	148.00	275.48	5.6e-08	528	abrin-a precursor - Indian lioc
P1R2:SI3431	+	138.00	255.50	7.3e-07	538	abrin-d precursor - Indian lioc
P1R2:SI16022	+	133.00	254.94	7.4e-07	562	abrin-c precursor - Indian lioc
P1R2:SI3430	+	133.00	249.53	1.6e-06	527	abrin-b precursor - Indian lioc
P1R2:SE3627	+	121.50	221.86	5.1e-05	570	agglutinin I precursor - Europ
P1R2:BS3761	+	99.50	173.92	0.0110	106	agglutinin I precursor - I
P1R2:CG3961	+	93.50	173.28	0.0584	251	abrin (clone 7.2) precursor -
P1R1:RLP6G	+	92.00	158.12	0.0875	286	rRNA N-glycosidase (EC 3.2.2.22
P1R2:SI13353	+	87.00	169.17	0.3171	300	N-acetylglucosaminidase (EC 3.2.2.22
P1R2:SI28821	+	84.00	152.89	0.6817	294	rRNA N-glycosidase (EC 3.2.2.22
P1R2:CG87585	-	79.00	145.83	2.34	212	hypothetical protein CC2714 [hm
P1R2:OE0401	-	79.00	143.97	2.41	261	antiviral protein - Virgintan R
P1R2:CG5032	+	78.00	142.47	3.09	247	karasurin B - Trichosanthes kid
P1R2:CG5606	+	76.50	138.06	4.64	289	karasurin C - Trichosanthes kid
P1R2:SA48020	+	76.00	138.47	5.16	247	karasurin B - Mongolian snake-gou
P1R2:SA2494	+	74.00	133.45	8.76	277	kinasin-related protein katc -
P1R2:SI17757	+	74.00	132.36	8.92	313	rRNA N-glycosidase (EC 3.2.2.22
P1R2:TA3963	+	73.50	129.71	10.12	730	translation elongation factor E
P1R2:TA02842	+	73.50	122.71	11.70	822	hypothetical protein IL231.2 [1
P1R2:TA13177	+	73.00	113.62	13.77	1038	soy protein - fruit fly (Dros
P1R2:TA1802	+	72.50	113.89	11.55	135	hypothetical protein F35612.9 -
P1R2:HA4355	-	72.50	130.75	12.79	268	hypothetical protein Vag2059h [1
P1R2:RTM7	-	72.50	130.07	12.94	289	rRNA N-glycosidase (EC 3.2.2.22
P1R2:TA7469	+	71.50	113.83	15.19	152	hypothetical protein SC1A11.02c
P1R2:TA67247	+	71.00	117.99	22.10	797	probable secreted protein. [lmg
P1R2:TA44066	+	70.00	116.78	28.18	730	translation elongation factor E
P1R2:TA61221	+	69.50	112.49	33.84	1054	probable calcium transporter 1
P1R1:JUD0146	-	69.50	112.60	34.74	533	serine proteinase (EC 3.4.21.-)
P1R2:BG9071	-	69.00	114.27	36.72	768	protein containing Cbm-repeats
P1R2:SI03135	-	68.50	122.76	35.65	773	actin-capping protein alpha cha
P1R2:JT07035	+	68.50	121.28	36.53	416	rRNA N-glycosidase (EC 3.2.2.22
P1R2:TA31256	+	68.50	118.12	38.51	350	terminal oxygenase component 1d
P1R2:SI23821	+	68.00	122.33	40.12	251	hypothetical protein 1 - feline
P1R1:ASLJF	+	68.00	121.93	40.40	263	vif protein - feline immunodef
P1R2:SI59430	+	66.00	116.79	44.01	467	hypothetical protein YD8223W -
P1R1:SI26191	-	66.00	116.67	44.10	473	nitrogenase (EC 1.18.6.1) molyp

plR2:AA5627	+	68.00	103.95	54.49	1957	myosin heavy chain isoform 1
plR2:JUN0108	+	67.50	121.39	45.58	250	luffin-b - smooth loofah
plR2:SS2519	+	67.50	120.44	46.31	278	beta-luffin - smooth loofah
plR2:AS2519	-	67.50	118.85	47.55	332	molB protein homolog - Rhodospirillum rubrum
plR2:157140	-	67.50	112.14	53.16	702	DNA translocase stage III sp
plR2:CG9999	+					

seq\_name: plr2:PD0018

```
seq_documentation_block:
  seq_documentation_block: Viscum album (fragment)
  seq_documentation_block: Viscum album (fragment)
```

C:Species: Viscum album  
C:Date: 03-Jul-1998 #sequence\_revision 10-Jul-1998 #text\_change 26-Aug-1999  
C:Accession: PD0018  
R:Schuenburg, S.; Krauspenhaar, R.; Mikhailov, A.; Stoeva, S.; Betzel, C.; Voelter, H.  
Biochem. Biophys. Res. Commun. 247, 367-372, 1998  
A:Title: Primary structure and molecular modeling of mistletoe lectin I from Viscum album  
A:Reference number: PD0018; M0ID:98308123  
A:Accession: PD0018  
A:Molecule type: protein  
A:Residues: 1-254 <ESCS>  
C:Superfamily: ricin; rRNA N-glycosidase homology  
F:1-246/Domain: rRNA N-glycosidase homology <RNG>

alignment_scores:	
Quality:	317.00
Ratio:	4.342
Percent Similarity:	82.955
	Length: 88
	Gaps: 0
	Percent Identity: 71.599

alignment\_block: PD0018

Align seg 1/1 to: PD0018 from: 1 to: 254

[illegible]

seq\_name: p1r1:RLCSD

```

seq_documentation_block:
  ricin D precursor - castor bean
N:Contains: RNA N-glycosidase (EC 3.2.2.22)
C:Species: Ricinus communis (castor bean)
C:date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903
R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.F.
Nucleic Acids Res. 13, 8019-8033, 1985
A:title: Genomic cloning and characterization of a ricin gene from Ricinus communis
A:Reference number: A24041; M01D:86067214
A:Accession: A24041
A:molecule type: DNA
A:residues: 1-576 <HML>

```



F:35-281/Domain: rRNA N-glycosidase homology <RNG>  
 F:303-564/Product: agglutinin chain B #status experimental <BCH>  
 F:319-361,362-402,405-443,450-485,489-528,531-564/Region: 40-residue repeats  
 F:324,259/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:304,147,231,233/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:200,203/Active site: Glu, Arg #status predicted  
 F:382-306,322-341,365-382,453-466,492-509/Disulfide bonds: #status predicted  
 F:324,337,348/Binding site: N-acetylglucosamine (Asp, Glu, Asn) #status predicted  
 F:397,437/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:536,557/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

## alignment\_scores:

Quality: 177.00 Length: 119  
 Ratio: 2.329 Gaps: 1  
 Percent Similarity: 63.866 Percent Identity: 36.975

## alignment\_block:

US-09-627-165B-13 x RLCSAG ..

Align seg 1/1 to: RLCSAG from: 1 to: 564

```

1  GCCAGATTCATCCCATCNCNTGTGAGCGTTCGCCGCAATTACAGTGG 50
   |||||.....||| .....||| .....||| .....|||
202  AAlArpPheGlnITyrlEgluGluMetArGThrArGlleArGTYras 218
   :::: ||| ||||| ::::: |||||:::|||||
218  nArGArSerAlaProAspProSerValIleThrLeuGlnuSnerITpg 235
   :::: ||| ||||| ::::: |||||:::|||||
101  GTGCAATTCACACCCAGATTCAGAGTGTCAAGATGCGATTATTAATAC 150
   ||||| .....||| .....||| .....||| .....|||
235  IYArGleuSerThrAlaIleGlnuSerAraGlnGlyAlaPheAlaSer 251
   ||||| .....||| .....||| .....||| .....|||
151  CAATTAAGATTCAGATTCGCCGCGTACCTTGTGACGNTGACAGTGGT 200
   ||||| .....||| .....||| .....||| .....|||
252  ProlIeGlnuLeuGlnArGArGAsnIlySerIySphaAsnValTYrasPva 268
   ||||| .....||| .....||| .....||| .....|||
201  TCGCAGCTGATCTCCACCTTGGCGATCATGTTGTTCGAATGACAGTGGT 250
   |.....| .....||| .....||| .....||| .....|||
268  IserIleuIleuIleProIleIleAlaIleuMetValTYrArGysAlaProP 285
   ||||| .....||| .....||| .....||| .....|||
251  GGCATTTCTCTCTGTGACACACCTTGGCGGTCTCTTAAGTCCGTC 300
   ||| |||.....||| .....||| .....||| .....|||
285  rOPrSerSerGlnPheSer.....LeuIleuIleArGProVal 297
   ||||| .....||| .....||| .....||| .....|||
301  GTGATGGCGGCAACGATGTCACTGTCGACTTTTCGCAACCCACCGTCG 350
   ||||| .....||| .....||| .....||| .....|||
298  ValPProAsnPheAsnAlaAspValIlySmetLasProGluProIleValAr 314
   ||||| .....||| .....||| .....||| .....|||
351  CATCGTA 357
   |||||
314  gtlleVal 316

```

seq\_name: p1r1:TLISA

seq\_documentation\_block:

abrin-a precursor - Indian licorice (fragment)  
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C:Species: Abrus precatorius (Indian licorice)  
 C:Date: 31-Dec-1993 #sequence, revision 01-Aug-1997 #text, change 16-Jul-1999  
 C:Accession: S32429; J70202; A39761; JCI398; S14472; S24133; S74110; S74111  
 R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
 J. Mol. Biol. 229, 263-267, 1993  
 A:Title: Primary structure of three distinct isoabrin determined by cDNA sequencing. CD  
 A:Reference number: S32429; MUID:93132798  
 A:Accession: S32429  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 'E', 2-528 <HUN>  
 A:Cross-references: GB:M98344; NID:q166294; PIDN:AAA3624.1; PID:q166295  
 A:Note: the coding region for the sequence shown is preceded by an ATG codon  
 A:Note: residues 1-8 were derived from the synthesized primer

R:Funatsu, G.; Taguchi, Y.; Kamenosono, M.; Yanaka, M.

Agric. Biol. Chem. 52, 1095-1097, 1988

A:Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein

A:Reference number: J70202

A:Accession: J70202

A:Molecule type: protein

A:Residues: 1-201,203-251 <FUN>

A:Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we h

R:Evensen, G.; Mathiesen, A.; Sundan, A.

J. Biol. Chem. 266, 6848-6852, 1991

A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.

A:Reference number: A39761; MUID:91201329

A:Accession: A39761

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 'E', 2-251 <EVE>

A:Cross-references: GB:X54872

A:Note: residues 1-8 were derived from the synthesized primer

R:Kimura, M.; Sumizawa, Y.; Funatsu, G.

Biosci. Biotechnol. Biochem. 57, 166-169, 1993

A:Title: The complete amino acid sequences of the B-chains of abrin-a and abrin-b, to

A:Reference number: JCI398; MUID:93169023

A:Contents: seeds

A:Accession: JCI398

A:Molecule type: protein

A:Residues: 261-347, 'T', 349-351, 'A', 353-357, 'L', 359-528 <KIM>

A:Experimental source: seed

R:Evensen, G.; Mathiesen, A.; Sundan, A.

submitted to the EMBL Data Library, October 1990

A:Description: Direct molecular cloning of two distinct abrin A-chains.

A:Reference number: S14471

A:Accession: S14472

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 'M', 2-251 <EV2>

A:Cross-references: EMBL:X54873; NID:q16090; PIDN:CAA3655.1; PID:q16091

R:Chen, Y.L.; Chow, L.P.; Tsugita, A.; Lin, J.Y.

FEBS Lett. 309, 115-118, 1992

A:Title: The complete primary structure of abrin-a B chain.

A:Reference number: S24133; MUID:92371656

A:Accession: S24133

A:Molecule type: protein

A:Residues: 262-297, 'Y', 299-426, 'L', 428-466, 'P', 468-482, 'L', 484-528 <CHE>

R:Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.

Eur. J. Biochem. 240, 564-569, 1996

A:Title: Probing the domain structure of abrin-a by tryptic digestion.

A:Reference number: S74110; MUID:97008945

A:Accession: S74110

A:Molecule type: protein

A:Residues: 89-108,154-172 <LIN>

A:Experimental source: seed

A:Accession: S74111

A:Molecule type: protein

A:Residues: 262-276, 'X', 278-280,329-348,369-388,399-418 <LIN>

A:Experimental source: seed

C:Comment: Abirin-a is more toxic than ricin. The toxin consists of an A chain, which

taining receptors on the cell surface. The A and B chains are linked by a single disu

C:Superfamily: ricin; rRNA N-glycosidase homology

C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglyutamic a

F:1-251/Product: abrin-a chain A #status experimental <ACH>

F:1-246/Domain: rRNA N-glycosidase homology <RNG>

F:261-528/Product: abrin-a chain B #status experimental <BCH>

F:283-325,326-366,369-407,414-449,453-492,495-528/Region: 40-residue repeats

F:2/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted

F:164,167/Active site: Glu, Arg #status predicted

F:247-269,286-305,329-346,417-430,456-473/Disulfide bonds: #status predicted

F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

F:361,401/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

alignment\_scores:

```

Align seg 1/1 to: T2LSA from: 1 to: 528

      1 GCGAGATTCAATCCCATCTCMTGTGAGAGCTTCGCCGGCAAATTACACTGG 50
          |||||||::: ||::: ||:::|||
      166 AAlaApPharIgtYrILeserAsnArgValArgValSerILleGIInthrGI 182
          |::: ||| ||:::|
      51 GGAGTCTNTCTCCACCACAACATGATGCTGCAGTGCGAGTGGAGACAGTTGGG 100
          |::: ||| ||:::|
      182 yThrAlAPheGlInProAsPAlalameTLleserLeuGluAsnASTTPA 199
          |::: ||| ||:::|
      101 GTGCGACATTCACACCCAAAGTCACAGACTCCAAGAGTGGAATTTTAATACC 150
          |::: ||| ||:::|
      199 sPaSnLseArIgrglValgIngInsVerValgInsPhrPheProAsn 215
          |::: ||| ||:::|
      151 CAAATAAGAATGTCAGATTTCCGCCGGTAACTTTGTACANTGACCAATGT 200
          ||||| ::|||
      216 GlH.....ValThLeuthrasnII 222
          .....
      201 TCGCAGAC.....GTGANTCTCA 217
          :||:::
      222 eArGsnglUpRoValILeValAsPserLseuSerHisPProThrValAlAY 239
          .....
      218 GGTGGCGCATGATGTTGTGCAATGACGATGGTGGCGGATTCCTCCCTCTC 267
          |||||::|
      239 alLeuAlaleuMetleuPheValCysAsn.....ProPrasnaIA... 252
          |||||::|
      268 GACCAACCTTCGCGCGTCTGCTCCTTAAGTCCGTGTGGATGCGGCCAACGA 317
          :::: |||||||:::|
      253 ...AngInsSerProLeuLeuIleArgSerILEValGLuLyseryls... 267
          :::: |||||||:::|
      318 TGTCACCTGCACCTNTTCCGAAACCCACCGGTGGCATC 354
          :::: |||||||
      268 .ILecySerSerArgTYrGIuPProThrValArgyle 279
          :::: |||||||

seq_name: pIr2:S32431

seq_documentation_block:
abrin-d precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian Licorice)
C>Date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 01-Aug-1997
C:Accession: S32431; S34408
R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J. Mol. Biol. 229, 263-267, 1993
A:Title: Primary structure of three distinct isoabryns determined by cDNA sequencing. CC
A:Reference number: S32429; MUID:93132798
A:Accession: S32431
A:Molecule type: mRNA
A:Residues: 1-528 <HDN>
A:Cross-references: GB:M98346
R:Hung, C.; Lee, M.; Lee, T.; Lin, J.
submitted to the EMBL Data Library, March 1993
A:Reference number: S34408
A:Accession: S34408
A:Molecule type: mRNA
A:Residues: 1-169, 'C', 171-320, 'L', 322-528 <HU2>
A:Cross-references: GB:M98346
C:Comment: Abryn consists of an A chain, which inhibits protein synthesis by inactivating
the A and B chains are linked by a single disulfide bond, which is essential for toxic
C:Superfamily: rich; RNA N-glycosidase homology
C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; P
F:1-251/Product: abrin-d chain A #status predicted <ACH>
F:1-246/Domains: rRNA N-glycosidase homology <RMG>
F:261-528/Product: abrin-d chain B #status predicted <BCH>
F:283-325-326-366-407-414-449-453-492-495-528/Region: 40-residue repeats
F:/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted

```

[illegible]

```

101  GTGCGACATGCCACCAAGTCGACAGATCCAAAGCAATGTCATTTT...AAT 147
      ::::: ||| ||||| ||||| ||| :::: ||| |||
199  spasnleuserc|yglYalginolnlseryalginasphrhnepheproasn 215
      ::::: ||| ::::: ||| ::::: ||| ::::: |||
148  ACCCAATTAATATTCGACAGATTCCGCCGGTACTTTGTACAGNTAGCAA 197
      ::::: ||| ::::: ||| ::::: ||| ::::: |||
216  AsnValnIleleuserSerIleasnArgGlnproValValalspsarLe 232
      ::::: ||| ::::: ||| ::::: ||| ::::: |||
198  TGTTCGCGACATGCATTCACAGCTTGAGGCAGCATGTTGTGCAATGCAGTG 247
      ::::: ||| ::::: ||||| ::::: ||||| |||||
232  userHisprothrValAlaValleAlaIleuMetLeuPheValCysasn. 248
      ::::: ||| ::::: ||||| ::::: ||||| ::::: |||||
248  GTCGGCCATTTCTCTCTCTGTGCAGCACCCCTTGCGCGCTGCTCTAAGGTC 297
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
249  ....ProProAsnAla.....AsnGlnSerProLeuIleuArgSer 261
      ::::: ||| ::::: ||| ::::: ||| ::::: |||
298  GTCTGTGATGGCGGCAACAGATGTCACCTGTGACTTTTTCGAAACCCACCGT 347
      ::::: ||| ::::: ||| ::::: ||| ::::: |||
262  IleValAlginIuglnSerIys...IleCysSerSerArgIyIuIuProthIyVa 277
      ::::: ||| ::::: ||| ::::: ||| ::::: |||
348  GCCGATC 354
      ||||| |||
277  IArgIle 279

```

```

seq_name: p1r2.S16022

seq_documentation_block:
  abrin-c precursor - Indian licorice
  N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
  C:Species: Abrus precatorius (Indian licorice)
  C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
  C:Accession: S16022
  R:Wood, K.A.; Lord, J.M.; Wawrzynczak, E.J.; Platak, M.
  Eur. J. Biochem. 198, 723-732, 1991
  A:Title: Preproabrin: genomic cloning, characterisation and the expression of the A-chain
  A:Reference number: S16022, M0ID:9126957
  A:Accession: S16022
  A:Status: preliminary
  A:Molecule type: DNA
  A:Residues: 1-562 <W00>
  A:Cross-references: EMBL:X55667; NID:g16084; PIDD:CM39202.1; PID:g16085
  C:Comment: abrin consists of an A chain, which inhibits protein synthesis by inactivating
  the A and B chains are linked by a single disulfide bond, which is essential for toxin
  C:Superfamily: ricin; rRNA N-glycosidase homology
  C:Keywords: duplication; glycoprotein; glycosidase; hydrolyase; lectin; pyroglyutamic acid
  F:15-35/Products: abrin-c chain A #status predicted <ACH>
  F:11-280/Domains: rRNA N-glycosidase homology <RNG>

```

F:295-562/Product: abrin-c chain B #status predicted <BC>  
 F:317-359, 360-400, 403-441, 448-483, 487-526, 529-562/Region: 40-residue repeats  
 F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
 F:108, 147, 229, 230/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:198, 201/Active site: Glu, Arg #status predicted  
 F:234, 287, 395, 435, 436/Binding site: carbonyl (Asn) (covalent) #status predicted  
 F:281-303, 320-339, 363-380, 451-464, 490-507/Disulfide bonds: #status predicted  
 F:322, 346/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
 F:534, 555/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

alignment\_scores:                      length:                      119  
 quality:                      138.00                      Gaps:                      4  
 Ratio:                      1.792                      Percent Identity:                      36.134  
 Percent Similarity:                      64.706

alignment\_block:  
 US-09-627-165b-13 x S16022

Align seg 1/1 to: S16022 from: 1 to: 562

```

1 GCCAGATTCATCCATCCTGAGAGCTTCGCCGCAATTAACAGTGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
200 Alarqytrargtrrlyrllseersanrlyglvalserlleavgrhcl 216
51 GGAGTCCTCCACCAACATGATGCTGAGCTGAGAGAGTGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||
216 YthralpneqlnProaspProalmetleuSerleuGlnAsnATPA 233
101 GTCGACATCCACCAAGTCCAGCAGTCCAGATGGATTTT..AAT 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||
223 spasnleuserglyglvalglnlnservalglinspThpProasn 249
148 ACCCAATAAATGATGAGATTCGCGCGGTACTTGTGACGNTGCAA 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||
250 AsnValilleuserSerlleAsnlnrglProvalvalAspserle 266
198 TGTTCGACGATGATCTCCAGCTTCGCGCATCTGTTGATGAGTGG 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||
266 userhlsProthrValAlaValAlaValAlaMetleuPheValCysAsn 282
248 GTCGGCATTCCTCTCTCGACCACTTCGCGCTGCTCTCAAGTCC 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||
283 .....ProProsnAla.....AsnlnSerProleuLeuIleArgSer 295
298 GTCGTGATGCGGCCAACGATGCTACCTGCACTTTTCCGACACCGCT 347
|||||:|||||:|||||:|||||:|||||:|||||:|||||
296 llevalgluluserlys...IleCysSerSerArgTyrGluProthVa 311
348 GCGCATC 354
|||||
311 lArgile 313

```

seq\_name: p1r2:S32430

seq\_documentation\_block:  
 abrin-b precursor - Indian licorice (fragment)  
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C:Species: Adrus precatorius (Indian licorice)  
 C:Date: 30-Sep-1993 #sequence\_revision 01-Aug-1997 #text\_change 20-Aug-1999  
 C:Accession: S32430; JCI399  
 R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
 J.Mol. Biol. 229, 263-267, 1993  
 A:Title: Primary structure of three distinct isobritins determined by cDNA sequencing. C  
 A:Reference number: S32429; MUID:93132798  
 A:Accession: S32430  
 A:Molecule type: rRNA  
 A:Residues: 1-527 <HUN>  
 A:Cross-references: GB:M98345; NID:q166296; PTDN:AA32625.1; PID:q166297  
 R:Kimura, M.; Sumizawa, T.; Funatsu, G.  
 Biosci. Biotechnol. Biochem. 57, 166-169, 1993  
 A:Title: The complete amino acid sequences of the B-chains of Abirin-a and Abirin-b, toxic  
 A:Reference number: JCI398; MUID:93169023

A:Accession: JCI399  
 A:Molecule type: protein  
 A:Residues: 260-281, 'D', 283-290, 'N', 292-349, 'PQ', 352-377, 'N', 379-425, 'W', 427, 'D', 429-  
 A:Experimental source: seed  
 C:Superfamily: ricin: rRNA N-glycosidase homology  
 C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin  
 F:1-250/Product: abrin-b chain A #status predicted <KCH>  
 F:7-245/Domains: rRNA N-glycosidase homology <RNG>  
 F:260-527/Product: abrin-b chain B #status experimental <BC>  
 F:282-324, 325-365, 368-406, 413-448, 452-491, 494-527/Region: 40-residue repeats  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted  
 F:110, 136, 400/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:74, 113, 194, 195/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:163, 166/Active site: Glu, Arg #status predicted  
 F:246, 268, 285-304, 328-345, 416-429, 455-472/Disulfide bonds: #status predicted  
 F:287, 311/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
 F:499, 520/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

alignment\_scores:                      length:                      120  
 quality:                      135.00                      Gaps:                      5  
 Ratio:                      1.667                      Percent Identity:                      33.333  
 Percent Similarity:                      67.500

alignment\_block:  
 US-09-627-165b-13 x S32430

Align seg 1/1 to: S32430 from: 1 to: 527

```

1 GCCAGATTCATCCATCCTGAGAGCTTCGCCGCAATTAACAGTGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||
165 Alarqytrargtrrlyrllseersanrlyglvalserlleavgrhcl 181
51 GGAGTCCTCCACCAACATGATGCTGAGCTGAGAGTGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||
181 ntrhalpneqlnProaspProalmetleuSerleuGlnAsnATPA 198
101 GTCGACATCCACCAAGTCCAGCAGTCCAGATGGATTTTAATAC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||
198 spasnleuserglyglvalglnlnservalglinspThpProasn 214
151 CAAATAGATTCAGATTTCCCGCGGTAC.....TTGTGACAGTGG 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||
215 AlavalhrleuArg...SerValAsnAsnGlnProvalillevalAspSe 230
195 CAATGTCGCGAGTATCTCCAGCTTCGCGCATCTGTTGCAATCA 244
|||||:|||||:|||||:|||||:|||||:|||||:|||||
230 rleuthrlhlslnservalAlaValAlaValAlaMetleuPheValCysA 247
245 GTCGTGACATTCCTCTCTCGACCACTTCGCGCTGCTCTTAAGG 294
|||||:|||||:|||||:|||||:|||||:|||||:|||||
247 sn.....ProProsnAla.....AsnlnSerProleuLeuIleArg 259
295 TCCGTCGTGATGCGGCCAACGATGCTACCTGCACTTTTCCGACACCG 344
|||||:|||||:|||||:|||||:|||||:|||||:|||||
260 Serillevalgluluserlys...IleCysSerSerArgTyrGluProth 275
345 CCGTCGATC 354
|||||
275 rValArgile 278

```

seq\_name: p1r2:S62627

seq\_documentation\_block:  
 agglutinin I precursor - European elder  
 C:Species: Sambucus nigra (European elder)  
 C:Date: 24-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 20-Aug-1999  
 C:Accession: S62627; S62619  
 R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.  
 Eur. J. Biochem. 235, 128-137, 1996  
 A:Title: The NeuAc(alpha2,6)-Gal/GalNAc-binding lectin from elderberry (Sambucus nig  
 A:Reference number: S62619; MUID:96202926  
 A:Accession: S62627







seq\_documentation\_block:  
 216 GGAGATCAGTCGCGACATTCGTCACNCATCGACAAAGTTACCGGGGAAA 167  
 ||||| .....  
 175 GYASpPhrLeuSerGIuIleAlaValAlaValThrGlyValSerMetAlaLy 191  
 .....  
 166 MCGCAATCTTAATTTGGTATTAATAAATGCC..... 136  
 ||||| .....  
 191 sLeuSInAlaTyraSngIyIleTyraSnaAlaAlaSnlYsIlleThrValGIyG 208  
 .....  
 135 .. ATCCCTTGAGATCGTCGACCTGGATTCGTCGACACCCCACTGCTTC 88  
 .....  
 208 InValLeuLysLeuThrGlyAlaIleSerSerIysProSerSerSer 224  
 .....  
 87 CAGCTCGAGCATGTACATGTT.....TGCTGAGANAGCT 53  
 .....  
 225 GlyLysIysTyValTyLeuProAlaSerAlaAspSerTrpArgIleTy 241  
 .....  
 52 CCCCACTGTTAATTTGCCGGCGAAG 28  
 ||||| ..  
 241 rProThrAsnLysAlaProValLys 249

seq\_name: p1r2:S28421

seq\_documentation\_block:  
 rRNA N-glycosidase (EC 3.2.2.22) PAP-alpha - Virginian pokeweed  
 N:Alternate names: antiviral protein alpha-pap  
 C:Species: Phytolacca americana (Virginian pokeweed)  
 C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 20-Jun-2000  
 C:Accession: S28421  
 R:Katooka, J.; Habuka, N.; Masuta, C.; Miyano, M.; Koiwai, A.  
 Plant Mol. Biol. 20, 879-886, 1992  
 A:Title: Isolation and analysis of a genomic clone encoding a pokeweed antiviral protein  
 A:Reference number: S28421, MIMD:93099240  
 A:Accession: S28421  
 A:Molecule type: DNA  
 A:Residues: 1294 <KAT>  
 A:Cross-references: EMBL:D10600; NID:g218010; PIDD:BA01451.1; PID:g218011  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 C:Keywords: glycosidase; hydrolase  
 F:30-278/Domain: rRNA N-glycosidase homology <RMS>

alignment\_scores:  
 Quality: 84.00 Length: 79  
 Ratio: 1.585 Gaps: 1  
 Percent Similarity: 67.089 Percent Identity: 26.582

alignment\_block:  
 US-09-627-165b-13 x S28421 ..

Align seg 1/1 to: S28421 from: 1 to: 294

1 GCCAGATTCATCCCATCTGTCGAGGCTTCGCCCGCAATTAACAGTGG 50  
 ||||| .....  
 201 AlaArgPheIysTyrIleGluAsnGlyValIleThrAsn..... 215  
 .....  
 51 GGAGTNTCTCCCAACATGTCATAGTCGAGCTGGAGACAGCTGGG 100  
 .....  
 216 .ArgAlaPheTyrProAsnAlaLysValIleAsnIleGluSerTrpG 232  
 ||||| .....  
 232 IyysIleSerThrAlaIleHisAsnAlaLysAsnGlyAlaLeuThrSer 248  
 .....  
 101 GTCCGACATCCACCCAGTCAGCAGCTCCAAAGATGTCATTTTAATACC 150  
 ||||| .....  
 249 ProLeuGluLeuLysAsnAlaAsnGlySerIysTrpIleValLeuArgVa 265  
 .....  
 201 TCCGCACTGATCTCCAGCTGGCGATCAGTGTTC 237  
 ||||| .....  
 265 IAspAspIleGluProAspValGlyLeuLeuLysTyr 277  
 .....  
 seq\_name: p1r2:C87585

seq\_documentation\_block:  
 hypothetical protein CC2714 [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: C87585  
 R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
 proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete genome sequence of Caulobacter crescentus.  
 A:Reference number: A87249; MIMD:21173698; PMID:11259647  
 A:Accession: C87585  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1212 <STO>  
 A:Cross-references: GB:AF005673; NID:g13424303; PIDD:AAK24679.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC2714

alignment\_scores:  
 Quality: 79.00 Length: 93  
 Ratio: 2.026 Gaps: 5  
 Percent Similarity: 41.935 Percent Identity: 27.957

alignment\_block:  
 US-09-627-165b-13/rev x C87585 ..

Align seg 1/1 to: C87585 from: 1 to: 212

259 AGATGCGCGACACCATTCGATTCAGCAACATGATGCCACGTCGAGATC 210  
 ||||| .....  
 97 ArgProAlaProHisCysProSerThrGlySerGlnGlnSerTyrAlaGly 113  
 .....  
 209 ACGTCGGGACATTCGTCACACATGTCACAAAGTTACCGGGGAAATCTGCCA 160  
 ||||| .....  
 113 sArgArgArgArgGlySerAlaGly..... 120  
 .....  
 159 TCTTATTTGGGTATTAATAATGCCATCTTGC..... 128  
 .....  
 121 .....ArgCysHisGlyTrpProAlaLysThrSerSer 131  
 .....  
 127 .....ACTGCTGAGCTTGGGTGATTCGACCCC.. 98  
 ||||| .....  
 132 ArgAsnAlaAlaArgProThrGlnGlyArgAsnTrpMetAlaArgProTr 148  
 .....  
 97 .AACTCGTCTCCAGCTCGAGCATGTCATGTTGCTGGAGANAGCTCCG 49  
 ||||| .....  
 148 pLeuSerSerIleThrCysAlaCysCys..IleTrpTrpArg.....Pro 162  
 .....  
 48 ACTGTTAATTTGCCGGCGAAGCTTCCA 22  
 .....  
 163 TrpArgArgTrpProSerGlyAlaPro 171  
 .....  
 seq\_name: p1r2:JE0401

seq\_documentation\_block:  
 antiviral protein - Virginian pokeweed  
 C:Species: Phytolacca americana (Virginian pokeweed)  
 C:Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 07-May-1999  
 C:Accession: JE0401  
 R:Kung, S.S.; Kimura, M.; Funatsu, G.  
 Agric. Biol. Chem. 54, 3301-3318, 1990  
 A:Title: The complete amino acid sequence of antiviral protein from the seeds of poke  
 A:Reference number: JE0401; MIMD:91242096  
 A:Accession: JE0401  
 A:Molecule type: protein  
 A:Residues: 1261 <KUN>  
 A:Experimental source: seed  
 C:Comment: This protein prevents the replication of a number of plant viruses, and in  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 C:Keywords: dsulfide bond; glycoprotein  
 F:6-254/Domain: rRNA N-glycosidase homology <RMS>

alignment\_scores:

Quality:	79.00	Length:	59
Ratio:	2.026	Gaps:	1
Percent Similarity:	66.102	Percent Identity:	33.898

alignment\_block:  
US-09-627-165B-13 x JE0401 . .

Align seg 1/1 to: JE0401 from: 1 to: 261

[illegible]



OM of: US-09-627-165B-13 to: SwissProt\_40:\* out\_format : pfs  
Date: Jul 30, 2002 4:07 PM

About: Results were produced by the Gencore software, version 4.5.  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=frame+...n2p.model -DEV=xip  
-O=/cgn2.1/USPRO.spool/US09627165/runat.30072002.151140\_3327/app-query.fasta.1.1012  
-DB=SwissProt\_40 -OPMT=fastan -SUFFIX=rsip -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPT=0.000 -LOPEXT=0.000  
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPO=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=Blotsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pcr  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTMT=pfis  
-NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09627165\_ECGN1.1.69 -NCPD=6 -ICPD=3 -LONGLOG  
-DEV=TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPHY -WAIT -THREADS=1

## Search information block:

Query: US-09-627-165B-13  
Query length: 357  
Database: SwissProt\_40:\*  
Database sequences: 105224  
Database length: 38719550  
Search time (sec): 36.450000

```
score_list:
Sequence      Strid Orig      ZScore      EScore Len      Documentation
SwissProt_40:MLA_VISAL + 317.00 624.82 1.6e-27 254 | P81446 viscum album (europ
SwissProt_40:R1C1_RICCO + 178.00 337.43 7.4e-12 576 | P02879 ricinus communis (cas
SwissProt_40:R1C1_RICCO + 178.00 335.61 9.5e-12 564 | P06750 ricinus communis (cas
SwissProt_40:ABRA_ABRPR + 148.00 277.81 1.7e-08 528 | P11140 abrus precatorius (in
SwissProt_40:ABRA_ABRPR + 138.00 257.10 2.3e-07 562 | P28590 abrus precatorius (in
SwissProt_40:ABRA_ABRPR + 135.00 251.64 4.8e-07 527 | P06077 abrus precatorius (in
SwissProt_40:NIGB_SAMNI + 108.00 196.66 0.0005 563 | P3183 sambucus nigra (europ
SwissProt_40:R1P1_MOMCH + 92.00 170.60 0.0291 286 | P16094 momordica charantia (
SwissProt_40:R1P1_PHYAM + 84.00 154.24 0.2311 294 | P03464 phytoacca americana
SwissProt_40:R1P2_BRIDI + 83.50 153.61 0.2612 282 | P28184 bryonia dioica (red b
SwissProt_40:R1P5_PHYAM + 79.00 145.25 0.8243 261 | P23339 phytoacca americana
SwissProt_40:R1P5_PHYAM + 76.50 139.29 1.60 289 | P24478 trichosanthos kirilow
SwissProt_40:KARC_ARATH + 74.00 129.56 2.13 754 | P46875 arabidopsis thaliana
SwissProt_40:R1P4_LUCFY + 74.00 134.64 3.03 267 | P00045 luffa cylindrica (smc
SwissProt_40:R1P1_CUCFI + 74.00 134.35 3.05 266 | P09174 cucumis fligera, puta
SwissProt_40:R1P1_CUCFI + 74.00 133.53 3.09 313 | P10297 phytoacca americana
SwissProt_40:EF2_METBU + 74.00 125.83 4.56 730 | P03632 mechanococoides burc
SwissProt_40:SOG_DROME + 72.50 120.61 4.88 1038 | P24025 drosophila melanogast
SwissProt_40:EF2_METBU + 70.00 117.77 10.00 730 | P09389 trichosanthos kirilow
SwissProt_40:EF2_METBU + 69.50 113.43 12.09 1054 | P43125 drosophila melanogast
SwissProt_40:R1P1_BRIDI + 69.00 114.15 11.10 290 | P3185 bryonia dioica (red b
SwissProt_40:PEPC_ASPNG + 69.00 118.62 12.29 533 | P33295 aspergillus niger, su
SwissProt_40:CARA_YEAST + 68.50 123.86 12.47 368 | P28495 saccharomyces cerevis
SwissProt_40:R1P3_GELAM + 68.50 122.36 12.82 316 | P33186 gelonium multiflorum
SwissProt_40:ST31_MOUSE + 68.50 111.73 15.57 1018 | P099m1 mus musculus (mouse)
SwissProt_40:VIR_FIYPE + 68.00 123.45 14.04 251 | P16089 feline immunodeficien
SwissProt_40:NIFD_FRAL + 68.00 117.69 15.59 473 | P02452 frankia alni, nitroge
SwissProt_40:R1B3_LUCFY + 67.50 122.48 15.96 250 | P22851 luffa cylindrica (smc
SwissProt_40:R1P2_PHYBG + 67.50 108.36 20.66 1182 | P15594 trypanosoma brucei g
SwissProt_40:R1P1_TRIAN + 67.00 120.00 18.66 294 | P56626 trichosanthos angust
SwissProt_40:305B_RAT + 67.00 119.06 18.98 326 | P31210 rattus norvegicus (rat)
SwissProt_40:U125_TITVT + 67.00 113.95 20.84 572 | P23987 infectious iaryngotira
SwissProt_40:PD2_CANAL + 67.00 110.50 22.19 836 | P06003 candida albicans (yea
SwissProt_40:INTR_BUTFA + 66.50 116.45 22.24 389 | P09052 bufo marinus (giant to
SwissProt_40:EF2_HALHA + 66.50 109.74 28.08 728 | P14823 halobacterium halobium
SwissProt_40:FLGI_BUCAI + 66.50 114.84 28.58 372 | P57426 bucheera aphidicola (
SwissProt_40:Y918_HUMAN + 65.50 106.16 33.49 966 | P09499 homo sapiens (human)
SwissProt_40:IR51_MOUSE + 65.50 103.94 34.88 1233 | P35559 mus musculus (mouse)
SwissProt_40:VL3_REOYD + 65.50 103.70 35.04 1267 | P17378 reovirus (type 3 / s
```

```
SwissProt_40:VL3_REOYD + 65.50 103.70 35.04 1267 | P17376 reovirus (type 1 /
SwissProt_40:RT03_MARPO - 65.00 112.51 33.31 430 | P26865 marchantia polymor
SwissProt_40:YDGB_SCHPO + 65.00 105.05 38.18 977 | P010495 schizosaccharomyce
SwissProt_40:PKP4_HUMAN + 65.00 103.10 39.57 1211 | P09569 homo sapiens (hum
SwissProt_40:SUBD_BACLI - 64.50 115.61 35.17 274 | P00781 bacillus licheniflo
```

seq\_name: SwissProt\_40:MLA\_VISAL

seq\_documentation\_block: ID MLA\_VISAL STANDARD; PRT; 254 AA.

AC P81446:  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DI 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Beta-galactoside specific lectin I A chain (MLA) (MU-I A) (rRNA N-  
glycosidase) (EC 3.2.2.22).  
OS Viscum album (European mistletoe).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Santalales; Viscaceae; Viscum.  
OX NCBI\_TaxID=3972;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=SUBSP. ALBUM;  
RX MEDLINE=97134581; PubMed=980141;  
RA Huguet Soler M., Stoeva S., Schwaborn C., Wilhelm S., Stiefel T.,  
RA Voelter W.;  
RT \*Complete amino acid sequence of the A chain of mistletoe lectin I.\*;  
RL FEBS Lett. 399:153-157(1996).

CC - FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE  
CC INTO THE CELL OF THE A CHAIN. B CHAINS ARE ALSO RESPONSIBLE FOR  
CELL AGGLUTINATION (LECTIN ACTIVITY).  
CC - CATALYTIC ACTIVITY: Endohydrolays of the N-glycosidic bond at one  
specific adenosine on the 28S rRNA.  
CC - SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
CC - MISCELLANEOUS: TWO ISOFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA  
AND NON-GLYCOSYLATED FORM MLA'.  
CC - SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 2  
RIP SUBFAMILY.

DR HSSP: P11140; IABR.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
DR PROSITE: PS00275; SHIGA\_RICIN; FALSE\_NEG.  
KM Hydrolase; toxin; Repeat; Glycoprotein; Lectin.  
FT ACT\_SITE 165 165  
FT CARBOHYD 112 112 N-LINKED (GLCNAc. . .).  
FT VARIANT 15 15 E -> D (IN MLA').  
FT VARIANT 66 66 V -> I (IN MLA').  
FT VARIANT 112 112 N -> T (IN MLA').  
FT VARIANT 116 116 P -> T (IN MLA').  
FT VARIANT 133 134 DQ -> EE (IN MLA').  
FT VARIANT 140 140 F -> Y (IN MLA').  
FT VARIANT 144 144 F -> Y (IN MLA').  
FT VARIANT 151 151 T -> A (IN MLA').  
FT VARIANT 179 179 Y -> D (IN MLA').  
FT VARIANT 184 184 A -> E (IN MLA').  
FT VARIANT 190 190 V -> M (IN MLA').  
FT VARIANT 218 218 I -> F (IN MLA').  
FT VARIANT 223 224 PP -> ST (IN MLA').  
FT VARIANT 231 231 T -> S (IN MLA').  
FT VARIANT 235 235 D -> S (IN MLA').

SEQUENCE 254 AA: 28478 MW; 53BAF98D5E0FEE67 CRC64;  
SQ

alignment\_scores: Quality: 317.00 Length: 88  
Ratio: 4.342 Caps: 0  
Percent Similarity: 82.955 Percent Identity: 71.591

alignment\_block:

US-09-627-165B-13 X MLA\_VISAL

Align seg 1/1 to: MLA\_VISAL from: 1 to: 254

```

1  GCCAGATTCAATCCCATCTCTGTTGGAGCGCTGGCGGGCAAAATTAACAGTGG  50
167  AAtArgPheAsnProIleLeuThrptArgTyrArgGlnTyrIleAsnSerI  133
51  GGAAGTCNTCTCCACCAACATGATACATGCTCGAGCTGGAGACGAGTTGGG  100
183  yAlaSerPheLeuProAspValTyrMetLeuGlnIleuGlnIuThrSerTrpG  200
101  GTGCACAAATCCACCCCAAGTCACGACGCTGCACGAGATGGCATTTTAAATAC  150
200  LygGlnInsThrGlnValGlnHisSerThrAspGlyValAlaPheAsnAsn  216
151  CAAATTAAGATTGCAGATTTCGGCCGCGTAACATTGTGTGACGNTGACCAATGT  200
217  ProIleArgIleuAlaIleProProGlyLysnBheValThrLeuThrAsnVa  233
201  TCCGCGACGTGATCTCCAGTTGGCATATGTTGTGCAGATCGCGTGC  250
233  lArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysGlyIuA  250
251  GGCCGATTCTCCTCT  264
150  rgrProSerSerSer  254

```

seq\_name: SwissProt\_40:RICI\_RICCO

```
seq_documentation_block:
ID   RICI_RICCO             STANDARD;          PRT;          576 AA.
```

DT 21-JUL-1986 (Rel. 01, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Richn precursor [Contains: Ricin A chain (rRNA N-glycosidase)  
 (EC 3.2.2.22); Ricin B chain].  
 DE Ricinus communis (Castor bean).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
 NCBI\_TaxID=3988;  
 [1]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=86067214; PubMed=2999712;  
 RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,  
 RA Weaver N.F.;  
 RT "Genomic cloning and characterization of a ricin gene from Ricinus  
 RT communis.";  
 RL Nucleic Acids Res. 13:8019-8033(1985).  
 [2]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=92163016; PubMed=1371405;  
 RA Tregear J.W., Roberts L.M.;  
 RT The lectin gene family of Ricinus communis: cloning of a functional  
 RT ricin gene and three lectin pseudogenes.";  
 RL Plant Mol. Biol. 18:515-525(1992).  
 [3]  
 SEQUENCE OF 12-576 FROM N.A.  
 RX MEDLINE=85179479; PubMed=3833723;  
 RA Lamb A., Roberts L.M., Lord J.M.;  
 RT "Nucleotide sequence of cloned cDNA coding for preproticlin.";  
 RL Eur. J. Biochem. 148:265-270(1985).  
 [4]  
 SEQUENCE OF 36-302.  
 RP Yoshitake S., Funatsu G., Funatsu M.;  
 RT Isolation and sequences of peptic peptides, and the complete  
 RT sequence of the chain of ricin-D.";  
 RL Agric. Biol. Chem. 42:1267-1274(1978).  
 [5]  
 SEQUENCE OF 315-576.

RA Funatsu G., Kimura M., Funatsu M.;  
RT "Primary structure of Ala chain of ricin D.";  
RL Agric. Biol. Chem. 43:2221-2224(1979).  
RN [6]  
RN CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.  
RX KIMURA Y., Kusuoku H., Tada M., Takagi S., Funatsu G.;  
RA MEDLINE=90344223; PubMed=1368517;  
RT "Structural analyses of sugar chains from ricin A-chain variant.";  
RL Agric. Biol. Chem. 54:157-162(1990).  
RN [7]  
RN REVIEW.  
RX MEDLINE=21480122; PubMed=11595634;  
RA Olsnes S., Kozlov J.V.;  
RT "Ricin.";  
RL Toxicol. 39:1723-1728(2001).  
RN [8]  
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
RX MEDLINE=87165983; PubMed=3558397;  
RA Monfort M., Villalaz J.E., Monzingo A.F., Ernst S.R., Katzin B.,  
RT Rutenber E., Xuong N.H., Hamlin R., Robertus J.D.;  
RL "The three-dimensional structure of ricin at 2.8 A.";  
RN J. Biol. Chem. 262:5398-5403(1987).  
RN [9]  
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
RX MEDLINE=91352004; PubMed=1881881;  
RA Katzin B.J., Collins E.J., Robertus J.D.;  
RT "Structure of ricin A-chain at 2.5 A.";  
RL Proteins 10:251-259(1991).  
RN [10]  
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.  
RX MEDLINE=91352005; PubMed=1881882;  
RA Rutenber E., Robertus J.D.;  
RT "Structure of ricin B-chain at 2.5-A resolution.";  
RL Proteins 10:260-269(1991).  
RN [11]  
RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.  
RX MEDLINE=95082010; PubMed=7990130;  
RA Weston S.A., Tucker A.D., Thatcher D.R., Derlyshire D.J.,  
RT Pauplt R.A.;  
RL "X-ray structure of recombinant ricin A-chain at 1.8-A resolution.";  
RN J. Mol. Biol. 244:410-422(1994).  
RN [12]  
RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.  
RX MEDLINE=96374222; PubMed=8780513;  
RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,  
RT Molina-Svinyh M.C., Robertus J.D.;  
RL "Structure and activity of an active site substitution of ricin A chain.";  
RN Biochemistry 35:11098-11103(1996).  
RN [13]  
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
RX MEDLINE=97240820; PubMed=9086280;  
RA Yan X., Hollis T., Svinyh M., Day P., Monzingo A.F., Milne G.W.,  
RT Robertus J.D.;  
RL "Structure-based identification of a ricin inhibitor.";  
RN J. Mol. Biol. 266:1043-1049(1997).  
RN [14]  
RN MUTAGENESIS.  
RX MEDLINE=93165632; PubMed=1287657;  
RA Kin Y., Robertus J.D.;  
RT "Analysis of several key active site residues of ricin A chain by mutagenesis and X-ray crystallography.";  
RL Protein Eng. 5:775-779(1992).  
RN -I- FUNCTION: Ricin is highly toxic to animal cells and to a less extent to plant cells. The A chain is responsible for inhibiting protein synthesis through the catalytic inactivation of 60S ribosomal subunits. It act as a glycosidase that removes that a specific adenine residue from an exposed loop of 28S ribosomal RNA. As this loop is involved in the binding of elongation factors, the modified ribosomes are unable to support protein synthesis. The A chain can inactivate a few thousands ribosomes per minute, thus inactivating them faster than the cell can make new ones. A single A-chain molecule can therefore kill an animal



```

RA MEDLINE=86059449; PubMed=2999130;
RX "Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;
RT "The primary sequence of Ricinus communis agglutinin. Comparison with
RL "richin.";
RN J. Biol. Chem. 260:15682-15686(1985).
RP [2]
RP SEQUENCE OF 303-564.
RC Araki T., Yoshioke Y., Funatsu G.;
RA "The complete amino acid sequence of the B-chain of the Ricinus
RT communis agglutinin isolated from large-grain castor bean seeds.";
RN Biochim. Biophys. Acta 872:277-285(1986).
RP [3]
RP SEQUENCE OF 303-337.
RX MEDLINE=80178723; PubMed=6768555;
RA Lin T.T.-S., Li S.S.-L.;
RT "Purification and physicochemical properties of ricins and
RL agglutinins from Ricinus communis.";
RN Eur. J. Biochem. 105:453-459(1980).
CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -I- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
CC INACTIVATING FAMILY, TYPE 2 RIP SUBFAMILY.
CC -I- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
-----
CC
EMBL: M12089; AAA33869.1; -
DR EMBL: S40368; AAB2584.1; -
DR PIR: A24261; RLC5AG.
DR HSP: P02879; IBR6.
DR GLYCOSULEDB: P06750; -
DR InterPro: IPR001574; RIP.
DR InterPro: IPR000772; Ricin_B_lectin.
DR Pfam: PF00652; Ricin_B_lectin; 2.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS50231; RICIN_B_LLECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;
KW Lectin; Signal.
FT SIGNAL 1 24
FT CHAIN 25 290
FT PROPEP 291 302
FT CHAIN 303 564
FT DOMAIN 309 436
FT DOMAIN 439 563
FT REPEAT 319 361
FT REPEAT 362 402
FT REPEAT 405 437
FT REPEAT 450 485
FT REPEAT 489 528
FT REPEAT 531 558
FT ACT_SITE 200 200
FT DISULFID 282 306
FT DISULFID 322 341
FT DISULFID 365 382
FT DISULFID 453 466
FT DISULFID 492 509
FT CARBOHYD 34 34
FT CARBOHYD 259 259
FT CARBOHYD 397 397
FT CARBOHYD 437 437
FT CONFLICT 331 331
FT CONFLICT 362 362
FT CONFLICT 374 374
FT CONFLICT 374 374
R -> G (IN REF. 2).
R -> G (IN REF. 2).

```

[illegible]



RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic  
 RT protein from the seeds of *Abrys precatorius*."; [3]  
 RN Agric. Biol. Chem. 52:1095-1097(1988).  
 RN [3]  
 RP SEQUENCE OF 1-251 FROM N.A.  
 RC TISSUE-Leaf;  
 RX MEDLINE=91201329; PubMed=1506300;  
 RA Evensen G., Mathiesen A., Sundan A.;  
 RT "Direct molecular cloning and expression of two distinct abrin  
 RT A-chains."; [4]  
 RL J. Biol. Chem. 266:6848-6852(1991).  
 RN [4]  
 RP SEQUENCE OF 262-528.  
 RX MEDLINE=92371656; PubMed=1505674;  
 RA Chen Y.-L., Chow L.-P., Tsungita A., Lin J.-Y.;  
 RT "The complete primary structure of abrin-a B chain."; [5]  
 RL FEBS Lett. 309:115-118(1992).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).  
 RX MEDLINE=9533188; PubMed=7608980;  
 RA Tshirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;  
 RT "Crystal structure of abrin-a at 2.14 A."; [1]  
 RL J. Mol. Biol. 250:354-367(1995).  
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
 CC ABRIN-A IS MORE TOXIC THAN RICIN.  
 CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
 CC PRECEDES ENDOCYTOSIS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-  
 CC INACTIVATING FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.  
 CC -----  
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M98344; AAA32624.1; ALT. INT.  
 DR EMBL; X54872; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; S52429; TZLSA.  
 DR PIR; S24133; S24133.  
 DR PDB; 1ABR; 07-FEB-95.  
 DR InterPro: IPR001574; RIP.  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 2.  
 DR Pfam; PF00161; RIP; 1.  
 DR SMART; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KM Hydrolyase; protein synthesis inhibitor; toxin; Repeat; Glycoprotein;  
 KM Lectin; 3d-structure.  
 KW CHAIN 1 251 ABRIN-A A CHAIN.  
 FT PEPTIDE 252 261 LINKER PEPTIDE.  
 FT CHAIN 262 528 ABRIN-A B CHAIN.  
 FT DOMAIN 273 400 RICIN B-TYPE LECTIN 1.  
 FT DOMAIN 403 527 RICIN B-TYPE LECTIN 2.  
 FT REPEAT 283 325 1-ALPHA.  
 FT REPEAT 326 366 1-BETA.  
 FT REPEAT 367 401 1-GAMMA.  
 FT REPEAT 402 449 2-ALPHA.  
 FT REPEAT 450 492 2-BETA.  
 FT REPEAT 493 528 2-GAMMA.

FT ACT\_SITE 164 164 BY SIMILARITY.  
 FT DISULFID 247 269 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 286 305 BY SIMILARITY.  
 FT DISULFID 329 346 BY SIMILARITY.  
 FT DISULFID 417 430 BY SIMILARITY.  
 FT DISULFID 456 473 BY SIMILARITY.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .).  
 FT CONFLICT 202 202 MISSING (IN REF. 2).  
 FT CONFLICT 298 298 N -> Y (IN REF. 4).  
 FT CONFLICT 427 427 M -> L (IN REF. 4).  
 FT CONFLICT 467 467 T -> P (IN REF. 4).  
 FT CONFLICT 483 483 V -> L (IN REF. 4).  
 FT SEQUENCE 528 AA; 59243 MW; A1F76BECDB9A827 CRC64;  
 alignment\_scores: Length: 129  
 Quality: 148.00 Gaps: 5  
 Ratio: 1.922  
 Percent Similarity: 59.690 Percent Identity: 33.333  
 alignment\_block:  
 US-09-627-165b-13 x ABRA\_ABRPR ..  
 Align seg 1/1 to: ABRA\_ABRPR from: 1 to: 528  
 1 GCCAGATTCATCCCATGTCGAGGAGCTTCGCGGCAATTAACAGTGG 50  
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
 166 AAlArgrhPheArgrYrIleSerASArgrValArgrAlserIlegrInrIgl 182  
 51 GGAGCTGTCCTCCACCAACATGATCATGCTCGAGCTGGAGACAGATTGGC 100  
 ::::|:::|||||:::|||||:::|||||:::|||||:::|||||  
 182 YThrAlArPhegrInrProASrAlaAlaMetIleSerLeuGlnASrnrIrrPA 199  
 101 GTCCAGCATGCCACCCAGTCGACGAGCAGTGGCATTTTAATAC 150  
 ::::|:::|||||:::|||||:::|||||:::|||||:::|||||  
 199 SPASrLeuSerArgrGlyAlaGlnIuSerValGlnASrInrPheProASn 215  
 151 CAATTAAGATTCGAGATTTCCGCGGTAATCTTGACGNTGACCAATGT 200  
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||  
 216 Gln.....ValrThLeuThrASnI 222  
 201 TCGGCAC.....GTGATTCACA 217  
 222 eArgrASnIuPrOvAlIleValASrSerLeuSerHnrIarProThrValAlav 239  
 218 GCTTGGCAGTCATGTTGTCGAATGCAAGTGGCGGCATTTCTCTCTCTC 267  
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||  
 239 AlLeuAlaIleuMetLeuPheValCysASn.....ProPrASnAla... 252  
 268 GACACACCTTCGCGCGCTCTCTTAAGTCCGTGATCGGCGCAACGA 317  
 ::::|:::|||||:::|||||:::|||||:::|||||:::|||||  
 253 ...ASnGrInSerProLeuIleuIleArgrSerIleValGluIuSerIlys.. 267  
 318 TGTCACTGCACTNTTCCGAACCCACGCGTGCATC 354  
 ::::|:::|||||:::|||||:::|||||:::|||||:::|||||  
 268 .IleCysSerSerArgrGlyrGlnuProThrValArgrIle 279  
 seq\_name: swissprot\_40:ABRC\_ABRPR  
 seq\_documentation\_block:  
 ID ABRC\_ABRPR STANDARD; PRT; 562 AA.  
 AC P28590;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Abirin-c precursor [contains: Abirin-c A chain (rRNA N-glycosidase)  
 DE (EC 3.2.2.22); Abirin-c B chain].  
 OS Abrys precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreeae; Abrys.

OX NCBI\_TaxID=3816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Leaf;  
 RX MEDLINE=91266957; PubMed=2050149;  
 RA Wood R.A., Lord J.M., Mawiznczak E.J., Piatk M.;  
 RT "Preproabrin: genomic cloning, characterisation and the expression of  
 RL the A-chain in Escherichia coli."; Eur. J. Biochem. 198;723-732(1991).  
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE  
 CC B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE  
 CC BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-  
 CC INACTIVATING FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X55667; CA39202.1; -  
 DR PIR: S16022; S16022.  
 DR HSP: P1140; IABR.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 2.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PRO0396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 DR HydroLase: Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;  
 KW Lactin; Signal  
 FT STGNL 1 34  
 FT CHAIN 1 34  
 FT PEPTIDE 286 295  
 FT CHAIN 296 562  
 FT DOMAIN 307 434  
 FT DOMAIN 437 561  
 FT REPEAT 317 359  
 FT REPEAT 360 400  
 FT REPEAT 403 435  
 FT REPEAT 448 483  
 FT REPEAT 526 562  
 FT REPEAT 529 562  
 FT ACT SITE 198 198  
 FT DISULFID 281 303  
 FT DISULFID 320 339  
 FT DISULFID 363 380  
 FT DISULFID 451 464  
 FT DISULFID 490 507  
 FT MOD\_RES 35 35  
 FT CARBOHYD 234 234  
 FT CARBOHYD 395 395  
 FT CARBOHYD 435 435  
 SO SEQUENCE 562 AA; 62817 MW; 1F0A0BCD7DBA6278 CRC64;

alignment\_scores: Quality: 138.00 Length: 119  
 Ratio: 1.792 Gaps: 4

Percent Similarity: 64.706 Percent Identity: 36.134

alignment\_block:

US-09-627-165b-13 x ABRRC\_ABRPR ..

Align seg 1/1 to: ABRRC\_ABRPR from: 1 to: 562

1 GCCAGATTCATCCATCCTNTGTGAGGCTTCCGCGCAAAATTAACAGTGC 50  
 200 AAlaGyTArGyTyrTleSerAsnArgValGlyValSerIleArgThrG1 216  
 51 GGAGCNCCTCCACCAACATGATGATGCTGAGCTGAGAGAGTGG 100  
 216 YThAlaHehInProAspProAlaMetLeuSerLeuGlnuSnrTtP 233  
 101 GTGACATCCACCCAGTCGACGATGCAAGATGTCATTTT...AT 147  
 233 sPAsnLeuSerGlyGlyAlGlnInSerValGlnAspThrPheProAsn 249  
 148 ACCCAATAGATTCGAGATTTCCGCGGATACCTTGTGACAGTGA 197  
 250 AsnValIleuSerSerIleAsnArgGlnProValValAlaSpSerLe 266  
 198 TGTTCGCGAGCTGATCTCCAGCTTGCGGATCATGTTGTGCAATG 247  
 266 uSerHisProThrValAlaValIleuAlaMetLeuPheValCysAsn. 282  
 248 GTGCGCATTCCTCTCTGCGACCCCTTGCGGCTGCTCTTAAGTCC 297  
 283 .....ProProAsnAla.....AsnInSerProLeuIleArgSer 295  
 298 GTGCGATGGCGGCCAGCATGTCACCTGCACCTGACCTTTCCAGCC 347  
 296 IleValGlnGlnSerLys...IleCysSerSerArgTyrGlnProThr 311  
 348 GCGCATC 354  
 311 LArgIle 313

seq\_name: SwissProt\_40:ABRRC\_ABRPR

seq\_documentation\_block: ID ABRRC\_ABRPR STANDARD; PRT; 527 AA.  
 AC 006077; P81374;

DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE ABrin-b precursor [contains: ABrin-b A chain (rRNA N-glycosidase)  
 DE (EC 3.2.2.22); ABrin-b B chain].  
 OS ABrus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eustosids I; Fabales; Fabaceae; Papilionoideae; Abrace; Abrus.  
 OX NCBI\_TaxID=3816;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=93132798; PubMed=8421313;  
 RA Hung C.-H., Lee M.-C., Lee T.-Y.;  
 RT "Primary structure of three distinct isoabrin determined by cDNA  
 RT sequencing. Conservation and significance."; J. Mol. Biol. 229:263-267(1993).  
 RL [2]  
 RN SEQUENCE OF 260-527.  
 RC TISSUE=Seed;  
 RX MEDLINE=93169023; PubMed=7763422;  
 RA Kimura M., Sumitawa T., Funatsu G.;  
 RT "The complete amino acid sequences of the B-chains of abrin-a and  
 RT abrin-b, toxic proteins from the seeds of Abrus precatorius."; Biosci. Biotechnol. Biochem. 57:166-169(1993).  
 RL -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
 CC ABRIN-A IS MORE TOXIC THAN RICIN.

CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
 CC PRECEDES ENDOCYTOSIS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-  
 CC INACTIVATING FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).  
 CC -----  
 CC EMBL: M98345; AAA32625.1; -.  
 CC HSSP: P11140; 1ABR.  
 CC InterPro: IPR001574; RIP.  
 CC InterPro: IPR000772; Ricin\_B\_Lectin.  
 CC Pfam: PF00652; Ricin\_B\_Lectin; 2.  
 CC Pfam: PF00161; RIP; 1.  
 CC PRINTS: PR00396; SHIGARICIN.  
 CC SMART: SM00458; RICIN; 2.  
 CC PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 CC PROSITE: PS00275; SHIGA\_RICIN; 1.  
 CC K01 Hydrolyase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;  
 CC Lectin.  
 CC CHAIN 1 250 ABRIN-B A CHAIN.  
 CC PEPTIDE 251 260 LINKER PEPTIDE.  
 CC CHAIN 261 527 ABRIN-B B CHAIN.  
 CC DOMAIN 272 399 RICIN B-TYPE LECTIN 1.  
 CC REPEAT 402 526 RICIN B-TYPE LECTIN 2.  
 CC REPEAT 282 324 1-ALPHA.  
 CC REPEAT 325 365 1-BETA.  
 CC REPEAT 368 400 1-GAMMA.  
 CC REPEAT 413 448 2-ALPHA.  
 CC REPEAT 452 491 2-BETA.  
 CC REPEAT 494 527 2-GAMMA.  
 CC ACT\_SITE 163 163 BY SIMILARITY.  
 CC DISULFID 246 268 INTERCHAIN (BY SIMILARITY).  
 CC DISULFID 285 304 BY SIMILARITY.  
 CC DISULFID 328 345 BY SIMILARITY.  
 CC DISULFID 416 429 BY SIMILARITY.  
 CC MOD\_RES 455 472 PYROLIDONE CARBOXYLIC ACID (BY  
 CC SIMILARITY).  
 CC CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CONFLICT 282 282 N->D (IN REF. 2).  
 CC CONFLICT 291 291 D->N (IN REF. 2).  
 CC CONFLICT 350 351 AE->PO (IN REF. 2).  
 CC CONFLICT 378 378 S->N (IN REF. 2).  
 CC CONFLICT 426 426 L->M (IN REF. 2).  
 CC CONFLICT 428 428 Y->D (IN REF. 2).  
 CC CONFLICT 431 431 N->S (IN REF. 2).  
 CC CONFLICT 484 484 R->K (IN REF. 2).  
 CC CONFLICT 491 491 N->S (IN REF. 2).  
 CC CONFLICT 493 493 H->Y (IN REF. 2).  
 CC CONFLICT 502 502 R->G (IN REF. 2).  
 CC CONFLICT 509 509 E->Q (IN REF. 2).  
 CC CONFLICT 513 513 H->W (IN REF. 2).  
 CC CONFLICT 516 516 H->T (IN REF. 2).  
 CC SEQUENCE 527 AA; 59114 MW; 3253AB490CE9494A CRC64;

alignment\_scores: 135.00

Length: 120

Ratio: 1.667 Gaps: 5  
 Percent similarity: 67.500 Percent identity: 33.333  
 alignment\_block:  
 US-09-627-165b-13 x ABRB\_ABRPR ..  
 Align seq 1/1 to: ABRB\_ABRPR from: 1 to: 527  
 1 GCCAGATTCGATCCCAACNTGTGGAGCTTCGCCGCAATTAACAGTGG 50  
 165 AAlArgrtYrgrhelleSerYrArYvalGlyValSerIleArghrAs 181  
 51 GGAGTCNTCTCCACCAACATGTACATGCTCGACGCTGGAGACGAGTGG 100  
 181 ntrAlaPhglInPrbAspAlaAlaMetIleSerIleGluAsnTrpA 198  
 101 GTGACAAATCCACCAAGTCGACAGTCCAGAGTGGATTTTATATACC 150  
 198 spAnLeuSerGlyGlyAlaGlnGlnSerValGlnAspPhrProAsn 214  
 151 CAATATAGATTGACGATTTCCGCCGGTAAAC.....TTTGTGACGNTGAG 194  
 215 AAlvalrhrleuArY...SerValAsnAsnGlnProValIleValAsp 230  
 195 CAATGTTGCGAGACGTGATCTCCAGCTTGGCGATCATGTGTTGATGCA 244  
 230 rleuThrHisGlnSerValAlaValLeuAlaLeuMetLeuPhValCys 247  
 245 GTGGTGGCCCATCTCTCTGTCGACCAACCTTCGCCGCTGCTTAAG 294  
 247 sn.....ProProAsnAla.....AsnGlnSerProLeuLeuIleArY 259  
 295 TCCGTCGTGATGCGGCCCAACGATGTCACCTGCACTTTTCCGACCC 344  
 260 SerIleValGluYSerIleYs...IleCysSerSerArYgrIuProth 275  
 345 CGTGGCATC 354  
 275 rValArYgIle 278  
 seq\_name: SwissProt\_40:NIGB\_SAMNI  
 seq\_documentation\_block:  
 ID NIGB\_SAMNI STANDARD; PRT; 563 AA.  
 AC P33183; P33184; P93542;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-MAR-2002 (Rel. 41, last sequence update)  
 DT 01-MAR-2002 (Rel. 41, last annotation update)  
 DE Nigirin b precursor (agglutinin V) (SNMV) [contains: Nigirin b A chain  
 DE (rRNA N-glycosidase) (EC 3.2.2.22); Nigirin b B chain].  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids II; Adoxaceae; Sambucus.  
 ON NCBI\_TaxId:4202;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Bark;  
 RA MEDLINE-96215449; PubMed-8647092;  
 RA Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;  
 RT "Characterization and molecular cloning of Sambucus nigra agglutinin V  
 RT (nigirin b), a GalNC-specific type-2 ribosome-inactivating protein  
 RT from the bark of elderberry (Sambucus nigra).";  
 RL Eur. J. Biochem. 237:505-513(1996).  
 RN 121  
 RP SEQUENCE OF 26-49 AND 298-321.  
 RC TISSUE-Bark;  
 RX MEDLINE-9403077; PubMed-8406135;  
 RA Gibbs T., Gitoras L., Ferreras J.M., Rojo M.A., Iglesias R.,  
 RA Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;  
 RT "Isolation and partial characterization of nigirin b, a non-toxic  
 RT novel type 2 ribosome-inactivating protein from the bark of Sambucus  
 RT nigra L.";



RT characterization and use for preparation of immunotoxins.";  
 RL Eur. J. Biochem. 176:581-588(1988).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE=94356447; PubMed=8075985;  
 RA Ren J., Wang Y., Dong Y., Stuart D.I.;  
 RT "The N-glycosidase mechanism of ribosome-inactivating proteins  
 implied by crystal structures of alpha-momorcharin.";  
 RL Structure 2:7-16(1994).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).  
 RX MEDLINE=94192822; PubMed=8143869;  
 RA Husain J., Tickle I.J., Wood S.P.;  
 RT "Crystal structure of momordin, a type I ribosome inactivating  
 protein from the seeds of Momordica charantia.";  
 RL FEBS Lett. 342:154-158(1994).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=95344383; PubMed=7619070;  
 RA Huang O., Liu S., Tang Y., Jin S., Wang Y.;  
 RT "Studies on crystal structures, active-centre geometry and  
 dehydrating mechanism of two ribosome-inactivating proteins.";  
 RL Biochem. J. 309:285-298(1995).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1  
 CC RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X57682; CAA40869.1; -  
 DR PIR: S14273; RLPUHG.  
 DR PIR: S16490; S16490.  
 DR PDB: 1A8A; 22-JUN-94.  
 DR PDB: 1A8B; 22-JUN-94.  
 DR PDB: 1AHC; 22-JUN-94.  
 DR PDB: 1MGC; 31-MAY-94.  
 DR PDB: 1MRH; 07-FEB-95.  
 DR PDB: 1MRH; 07-FEB-95.  
 DR PDB: 1MRH; 07-FEB-95.  
 DR GlycoSuiteDB: P16094; -  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA-RICIN; 1.  
 DR Protein synthesis inhibitor; Hydrolase; Toxin; Signal; Glycoprotein;  
 KW 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 269 RIBOSOME-INACTIVATING PROTEIN MOMORDIN I.  
 FT PROPEP 270 286 MISSING IN MATURE PROTEIN.  
 FT ACT SITE 183 183  
 FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .).  
 FT /FTID-CAR\_000082.  
 FT SEQUENCE 286 AA; 31532 MW; E1B013ABEBC216CF CRC64;  
 SQ

alignment\_scores:  
 Quality: 92.00 Length: 82  
 Ratio: 1.769 Gaps: 3  
 Percent Similarity: 63.415 Percent Identity: 34.146

alignment\_block:

US-09-627-165b-13 x RIP1\_MOMCH ..  
 Align seg 1/1 to: RIP1\_MOMCH from: 1 to: 286

1 GCCAGATTCAATCCATCNCNTGTGACGCTTCCGCGCAATTAAAGATGG 50

seq\_name: SwissProt\_40:RIPA\_PHYAM  
 seq\_documentation\_block:  
 ID RIPA\_PHYAM STANDARD; PRT; 294 AA.  
 AC 003464; (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 01-OCT-1993 (Rel. 41, Last annotation update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Antiviral protein alpha precursor (PAP-alpha) (Ribosome-inactivating  
 DE protein) (rRNA N-glycosidase) (EC 3.2.2.22).  
 OS Physolacca americana (common pokeweed) (Virginia pokeweed).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Physolacaceae; Physolacca.  
 OC NCBI\_TaxID=3527;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seed, leaf, and root;  
 RX MEDLINE=93099240; PubMed=1281438;  
 RA Kataoka J., Habuka N., Masuta C., Miyano M., Koiwai A.;  
 RT "Isolation and analysis of a genomic clone encoding a pokeweed  
 RT antiviral protein.";  
 RL Plant Mol. Biol. 20:879-886(1992).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=95010127; PubMed=7925458;  
 RA Ago H., Kataoka J., Tsuge H., Hakuba N., Inagaki E., Noma M.,  
 RA Miyano M.;  
 RT "X-ray structure of a pokeweed antiviral protein, coded by a new  
 RT genomic clone, at 0.23-nm resolution. A model structure provides a  
 RT suitable electrostatic field for substrate binding.";  
 RL Eur. J. Biochem. 225:369-374(1994).  
 CC -1- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN  
 CC SYNTHESIS IN VITRO. HAS ALSO BEEN SHOWN TO INHIBIT THE  
 CC REPLICATION OF MAMMALIAN VIRUSES. THE PROTEIN MAY PROVIDE A  
 CC MEANS OF CELLULAR SUICIDE UPON INVASION BY A VIRUS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Cell wall.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1  
 CC RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D10600; BAA01451.1; -  
 DR PIR: S28421; S28421.  
 DR PDB: 1APA; 31-JAN-94.

cores:		
quality:	83.50	length:
ratio:	1.606	gaps:
		4

Percent Similarity: 62.651 Percent Identity: 31.325

alignment\_block:  
US-09-627-165B-13 x RIP2\_BRXDI ..

Align seg 1/1 to: RIP2\_BRXDI from: 1 to: 282

```

1 GCCACATTCATCCCATCCTGAGCGCTTCGCCGCAATTAACAGTGG 50
  |||||
165 AlaArgPheLysTyrIleGluGlnArgValSerGlu.....AsnValG1 199
51 GGAGTCNTCTCCACCAACATGTACATCTGACCTGAGACGAGTGGG 100
  |||||
199 yThrLysPheLysProAspProAlaPheLeuSerLeuGlnAsnAlaTrpG 216
101 GTTCGACATTCACCCAGTCCAG.....CAGTCCAGAGATGCATTTT 144
  |||||
216 LysLeuSerGluGlnIleGlnIleAlaGlnThrArgIleGlyGluPhe 232
146 AATACCAATAGATTCAGAT.....TCCGCGCTTAAGTTCGTGAC 188
  |||||
233 AlaArgProValGluLeuArgThrValSerAsnThrProThrPheValTh 249
189 GNTGACCATGTTCGCGACGTATCTCCAGCTTCGCGATCATCTGTTTC 237
  |||||
249 rAsnValAsn...SerProValValLysGlyIleAlaLeuLeuLeuTyr 264

```

seq\_name: SwissProt\_40:RIP2\_PHYAM

seq\_documentation\_block:

```

ID RIP2_PHYAM STANDARD; PRT: 261 AA.
AC P23339;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Activator protein S (PAP-S) (Ribosome-inactivating protein) (rRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Phytolacca americana (Common pokeweed) (Virginia pokeweed);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Phytolaccaceae; Phytolacca.
OX NCBI_TaxID=527;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed.
RA MEDLINE=91242096; PubMed=1368643;
RX Kung S.S., Kimura M., Funatsu G.;
RT "The complete amino acid sequence of activator protein from the seeds
RT of pokeweed (Phytolacca americana).";
RL Agric. Biol. Chem. 54:3301-3318(1990).
CC -1- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN
CC SYNTHESIS IN VITRO.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY, TYPE 1
CC RIP SUBFAMILY.
CC PIR: JEO401; JEO401.
DR HSSP: 003464; JAPA.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGARICIN.
KW Anti-viral; Protein synthesis inhibitor; Hydrolase; Toxin.
FT ACT_SITE 175 175
FT DISULFID 34 258
FT DISULFID 84 105
FT DISULFID 261 261
FT SEQUENCE 261 AA; 29200 MM; D88B99962FE8399D CRC64;

```

alignment\_scores:  
Quality: 79.00 Length: 59  
Ratio: 2.026 Gaps: 1  
Percent Similarity: 66.102 Percent Identity: 33.898

alignment\_block:

US-09-627-165B-13 x RIP2\_PHYAM ..

Align seg 1/1 to: RIP2\_PHYAM from: 1 to: 261

```

1 GCCACATTCATCCCATCCTGAGCGCTTCGCCGCAATTAACAGTGG 50
  |||||
177 AlaArgPheLysTyrIleGluGlnArgValSerGluThrAsnPheAsnArgAs 193
51 GGAGTCNTCTCCACCAACATGTACATCTGACCTGAGACGAGTGGG 100
  |||||
193 pHeSer.....ProAsnAspLysValLeuAspLeuGluAsnTrpG 208
101 GTTCGACATTCACCCAGTCCAGAGTCCAGAGATGCATTTTATATACC 150
  |||||
208 LysLeuSerThrAlaIleHisAsnSerLysAsnGlyAlaLeuProLys 224
151 CAATTAAGATTCAGATTCGCGCGCT 177
  |||||
225 ProLeuGluLeuLysAsnAlaAspGly 233

```

seq\_name: SwissProt\_40:RIP2\_TRIKI

seq\_documentation\_block:

```

ID RIP2_TRIKI STANDARD; PRT: 289 AA.
AC P24478;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein karasurin precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Trichosanthes kirilowii (Mongolian snake-gourd);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=3677;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root tuber;
RA MEDLINE=97356562; PubMed=9212998;
RX Mizukami H., Iida K., Kondo T., Ogihara Y.;
RT "Cloning and bacterial expression of a gene encoding ribosome-
RT inactivating proteins, karasurin-A and karasurin-C, from Trichosanthes
RT kirilowii var. japonica.";
RL Biol. Pharm. Bull. 20:711-713(1997).
RN [2]
RP SEQUENCE OF 24-270.
RA MEDLINE=92005921; PubMed=1914000;
RX Toyokawa S., Takeda T., Kato Y., Wakabayashi K., Ogihara Y.;
RT "The complete amino acid sequence of an abortifacient protein,
RT karasurin.";
RL Chem. Pharm. Bull. 39:1244-1249(1991).
CC -1- FUNCTION: ABORTION-INDUCING PROTEIN. IT INACTIVATES EUKARYOTIC
CC 60S RIBOSOMAL SUBUNIT.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY, TYPE 1
CC RIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AB000666; BAA21786.1; -
DR PIR: J00393; J00393.
DR HSSP: P09989; IMR1.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP; 1.

```

```

DR PRINTS: PR00396; SHICARICIN.
DR PROSITE: PS00275; SHICARICIN; 1.
KW Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin; Signal.
FT SIGNAL 1 21
FT CHAIN 22 270
FT PROPEP 24 270
FT ACT_SITE 271 289
FT ACT_SITE 183 183
FT SEQUENCE 289 AA; 31704 MW; 883D3E3242867B26 CRC64;
BY SIMILARITY.

alignment_scores:
  Quality: 76.50 Length: 99
  Ratio: 1.195 Gaps: 4
  Percent Similarity: 64.646 Percent Identity: 25.253

alignment_block:
US-09-627-165b-13 x RIPS_TRIKI ..
Align seg 1/1 to: RIPS_TRIKI from: 1 to: 289

1 GCCGAGTTCATCCATCTGTGGAGGCTTGGCGGCAATTAACAGTGG 50
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
185 AIAATGTYLTPHELLEGLUGLNGLNLGLYLSATGVALASP.... 199
51 GGAGTCTCTCCACCAACATGTACATGCTCGAGCTGGAGACGAGTTGG 100
  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
200 LYSTIRPHELAPROSERIEUAIETLEISERLEUGLNASISERTPS 216
101 GTGACCAATCCACCCAGTCCAG.....CAGTCCAAAGATGCGCATTTT 144
  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 ERALAEUSERTYSGINLEGLNLLEALASERTHRASNSNGLYNPHE 232
145 AATGCCAATAAGATGACATTTCCGCGGTAAGTCTTGAGCGTGG 194
  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 GIURHPROVALLEULEILASNLGINSGLNARGVALRHLRLEH 249
195 CAATGTCGCGAC.....GTGATCTCCAGCTTGCGGATCAGTTGTTG 238
  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 TASHVALAPALAGLYVALTHTSRASNLLEALAEULEULEU... 264
239 AATGACGTGTCGCCATTCCTCTCTCCACCCCTTCGCCGCGT 285
  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 .....ASNARGASNSMETALALALALASPSAPSPALPROMET 278

seq_name: SwissProt_40:KATC_ARATH

seq_documentation_block:
ID KATC_ARATH STANDARD; PRT; 754 AA.
AC P46875;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Kinesin-like protein C.
GN KATC.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_Taxid:3702;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RC MEDLINE=94355659; PubMed=8075402;
RA Mitsui H., Nakatani K., Yamaguchi-Shinozaki K., Shinozaki K.,
RA Nishikawa K., Takahashi H.;
RT "Sequencing and characterization of the kinesin-related genes katc
RT and katc of Arabidopsis thaliana."
RT Plant Mol. Biol. 25:865-876(1994).
CC -1- FUNCTION: POSSIBLE ROLE IN MITOSIS.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS; A SMALL GLOBULAR
CC N-TERMINAL, A CENTRAL ALPHA-HELICAL COILED COIL AND A LARGE
CC GLOBULAR C-TERMINAL WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY

```

```

CC (IT HYDROLYZES ATP AND BINDS MICROTUBULES).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D21138; BAA04674.1; -.
DR HSSP: P17119; 3KAR.
DR InterPro: IPR001752; kinesin.
DR Pfam: PF00225; kinesin.1.
DR PRINTS: PR00380; KINESINHEAVY.
DR SMART: SM00129; KISC; 1.
DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil;
KW Multigene family.
FT DOMAIN 1 45 GLOBULAR.
FT DOMAIN 46 393 COILED COIL.
FT DOMAIN 398 727 KINESIN-MOTOR (BY SIMILARITY).
FT NP_BIND 481 488 ATP (POTENTIAL).
FT SEQUENCE 754 AA; 85030 MW; 76091CD5B5D9C531 CRC64;

alignment_scores:
  Quality: 76.00 Length: 30
  Ratio: 3.167 Gaps: 0
  Percent Similarity: 80.000 Percent Identity: 50.000

alignment_block:
US-09-627-165b-13 x KATC_ARATH ..
Align seg 1/1 to: KATC_ARATH from: 1 to: 754

82 GAGCTGGAGACGAGTTGGGTGCAATCCACCAAGTCCACGACTCCAA 131
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 GHLIELGLTHTTTCYSERSESGINSERTHTGHLIEATGHLNLEUG 331
  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 GGATGCAATTTTAAATACCAATTAAGATTCAGATTTGCC 171
  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
331 NAAARGLEUVALANSERGLNARGLEUGLNVALSER 344

seq_name: SwissProt_40:RIPA_LUPCY

seq_documentation_block:
ID RIPA_LUPCY STANDARD; PRT; 277 AA.
AC Q00465;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein luffin-alpha precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22).
DE Luffa cylindrica (Smooth loofah) (Sponge gourd).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Cucurbitales; Cucurbitaceae; Luffa.
OX NCBI_Taxid:3670;
RN 111
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RC MEDLINE=92288316; PubMed=1600156;
RA Katsoka J., Habuka N., Miyano M., Masuta C., Koizumi A.;
RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-
RT inactivating protein from Luffa cylindrica."
RT Plant Mol. Biol. 18:1199-1202(1992).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1

```



RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: X62371; CAA44229.1; -  
 DR PTR; S22494; S22494.  
 DR HSSP; P16094; IARC.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin; signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 277  
 FT ACT\_SITE 179 179 RIBOSOME-INACTIVATING PROTEIN LUFCIN-  
 FT SEQUENCE 277 AA; 30212 MW; EAI7EC27998C25AC CRC64;  
 SQ

alignment\_scores:  
 Quality: 74.00 Length: 80  
 Ratio: 1.423 Gaps: 2  
 Percent Similarity: 65.000 Percent Identity: 26.250

alignment\_block:  
 US-09-627-165b-13 x RIP1\_LUFCY ..

Align seg 1/1 to: RIP1\_LUFCY from: 1 to: 277

```

1 GCCAGATTCAATCCATCTGTGAGGCTCCGCGCAATTACAGTGG 50
   ::::::::::::::::::::
181 SerArgPheLysTyrIleGlnGlnIleIleGlnArgIleSer..... 195
   ::::::::::::::::::::
51 GGAGCTCTCCACCAACATGTACATGCTGCAGCTGGAGCAGTGGG 100
   ::::::::::::::::::::
196 LysAsnGlnValProSerLeuAlaThrIleSerLeuGlnAsnGlnTrpS 212
   ::::::::::::::::::::
101 GTGACATATCACCACCAAGTCCAG.....CAGTCCAAAGCATGTTT 144
   ::::::::::::::::::::
212 erAlaLeuSerLysGlnIleGlnLeuAlaGlnThrAsnAsnGlnTrpPhe 228
   ::::::::::::::::::::
145 AATACCCAAATAGATTGCAGATTCCCGCGGTACTTGTGACGTTGAG 194
   ::::::::::::::::::::
229 LysThrProValValIleThrAspAspLysGlnArgValGlnIleThr 245
   ::::::::::::::::::::
195 CAATGTCGGGACGTGATCTCCACGCTGGGCGATCATGTTG 234
   ::::::::::::::::::::
245 AsnValThrSerLysValValThrLysAsnIleGlnLeu 258
   ::::::::::::::::::::

```

seq\_name: SwissProt\_40:RIP1\_CUCFI

seq\_documentation\_block:  
 ID RIP1\_CUCFI STANDARD; PRT; 286 AA.  
 AC Q9FRX4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE putative ribosome-inactivating protein precursor (rRNA N-glycosidase)  
 DE (EC 3.2.2.22).  
 OS Cucumis figareli.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.  
 OX NCBI\_TaxID=131071;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada T., Ohki S.T., Osaka T.;

RT "Cloning and analysis of a cDNA coding a putative ribosome-  
 RT inactivating protein from Cucumis figareli".  
 RL Plant Biotechnol. 17:337-340(2000).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1  
 CC RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: AB045560; BAB19677.1; -  
 DR InterPro: IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin; signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 286  
 FT ACT\_SITE 185 185 PUTATIVE RIBOSOME-INACTIVATING PROTEIN.  
 FT CARBOHYD 103 103 BY SIMILARITY.  
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 286 AA; 31771 MW; 4EPD4966E604DA41 CRC64;

alignment\_scores:  
 Quality: 74.00 Length: 62  
 Ratio: 1.762 Gaps: 1  
 Percent Similarity: 67.742 Percent Identity: 29.032

alignment\_block:  
 US-09-627-165b-13 x RIP1\_CUCFI ..

Align seg 1/1 to: RIP1\_CUCFI from: 1 to: 286

```

55 TCNTCTCCACCAACATGTACATGCTGCAGCTGGAGCAGTGGGGTGG 104
   ::::::::::::::::::::
203 SerLysValProAspLeuAlaIleSerLeuGlnAsnGlnTrpSerLe 219
   ::::::::::::::::::::
105 ACAATCCACCCCAAGTCCAG.....CAGTCCAAAGCATGTTTAAFA 148
   ::::::::::::::::::::
219 WLeuSerLysGlnIleGlnIleAlaLysSerAsnAsnGlnIlePheGlnT 236
   ::::::::::::::::::::
149 CCCAAATAGATTGCAGATTCCCGCGGTACTTGTGACGTTGACAAAT 198
   ::::::::::::::::::::
236 hrProValLysIleIleAsnAspLysGlnIleLeuThrGlnValThrAsn 252
   ::::::::::::::::::::
199 GTTCGCGAGCATGATCTCCACCTGTCGCGATCATGTTG 234
   ::::::::::::::::::::
253 ValSerSerLeuValValThrLysAsnIleMetLeu 264
   ::::::::::::::::::::

```





```

253 erThAspArgAlaSerSerAspGlnPheIleAspHis.....MetLeu 267
286 CTCCTAAGTCCGTCGTCG.....GATCGCGC 311
268 MetIleAlaGProIleLeuValAlaAspValAlaGluValAlaThrAspAlaAs 284
312 CAAGATGTCACCTGCACTTTTCCGAACCCAGCGTCGCATC 354
284 pAsnAspAspThrCysAlaAspProGluProThrValAlaArgIle 298

```

seq\_name: sp\_plant:Q94BW4

```

seq_documentation_block:
ID   Q94BW4      PRELIMINARY;   PRT;   580 AA.
AC   Q94BW4;
DT   01-DEC-2001 (TREMBLrel. 19, Created)
DT   01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE   TYPE 2 RIBOSOME-INACTIVATING PROTEIN CINNAMOMIN II
DE   PRECURSOR.
OS   Cinnamomum camphora (Camphor tree).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX   NCBI_TaxID=13429;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yang Q., Gong Z.Z., Liu W.Y.;
RT   "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT   genes encoding cinnamomil proteins and study of their expression
RT   patterns."
RL   Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AY039802; AAK82459.1; -.
KW   Signal.
FT   SIGNAL      1       32      POTENTIAL.
FT   CHAIN       33      580      TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT   SEQUENCE    580 AA; 64265 MW; 37E4289ECCECBPF CRC64;
SQ

```

#### alignment\_scores:

Quality: 179.00 Length: 131  
Ratio: 2.081 Gaps: 6  
Percent Similarity: 65.649 Percent Identity: 40.458

#### alignment\_block:

US-09-627-165B-13 x Q94BW4

Align seg 1/1 to: Q94BW4 from: 1 to: 580

```

4 AGATTCATCCCATCNCNTGTGAGGCTTCGCCGCAATTAACAGTGGGA 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
202 ArgPheArgPheIleGluTyrArgValArgGluSerIleThrArgAlaG 218
54 GTCNCTCCACCAACATGTACATGCTGAGCTGAGACGATGGGGTC 103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
218 UMetPheArgProAspProAlaMetLeuSerLeuGluAsnLysTrpSerA 235
104 GACAAATCCACCAAGTCACAGATCC...AAGATGSCATTTTAAATAC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
235 lAluSerAsnAlaValGlnGlnSerAsnLysProValTyrValGlySerVa 268
151 CAATTAAGATTGCAATTCGCCGCGTAACTTTGTGACGNTGACAAATG 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
252 ProValGluLeuArgSerIleSerAsnLysProValTyrValGlySerVa 268
201 TCGCGAC...GTGATCTCCAGCTTGGGATGATGTTGTTGATGATGC... 243
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 lSerAspArgValIleSerGlyLeuAlaIleMetLeuPheIleCysArgS 285
244 .AGTGGTCCGCAATTCCTCT.....CTCGACACCCCTTCGCGCGT 285
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
285 erSerAspArgAlaSerSerAspGlnPheIleAspHis.....LeuLeu 299

```

```

286 CTCCTAAGTCCGTCGTCG.....GATCGCGC 311
300 MetIleAlaGProIleLeuValAlaAspValAlaAspValAlaThrAspAlaAs 316
312 CAAGATGTCACCTGCACTTTTCCGAACCCAGCGTCGCATC 354
316 pAsnAspAspThrCysAlaAspProGluProThrValAlaArgIle 330

```

seq\_name: sp\_plant:Q94BW3

```

seq_documentation_block:
ID   Q94BW3      PRELIMINARY;   PRT;   580 AA.
AC   Q94BW3;
DT   01-DEC-2001 (TREMBLrel. 19, Created)
DT   01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE   TYPE 2 RIBOSOME-INACTIVATING PROTEIN CINNAMOMIN III
DE   PRECURSOR.
OS   Cinnamomum camphora (Camphor tree).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX   NCBI_TaxID=13429;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yang Q., Gong Z.Z., Liu W.Y.;
RT   "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT   genes encoding cinnamomil proteins and study of their expression
RT   patterns."
RL   Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AY039803; AAK82460.1; -.
KW   Signal.
FT   SIGNAL      1       32      POTENTIAL.
FT   CHAIN       33      580      TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT   SEQUENCE    580 AA; 64421 MW; 940D10F01EFB558 CRC64;
SQ

```

#### alignment\_scores:

Quality: 179.00 Length: 131  
Ratio: 2.106 Gaps: 5  
Percent Similarity: 64.885 Percent Identity: 38.931

#### alignment\_block:

US-09-627-165B-13 x Q94BW3

Align seg 1/1 to: Q94BW3 from: 1 to: 580

```

4 AGATTCATCCCATCNCNTGTGAGGCTTCGCCGCAATTAACAGTGGGA 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
202 ArgPheArgPheIleGluTyrArgValArgGluSerIleThrArgAlaG 218
54 GTCNCTCCACCAACATGTACATGCTGAGCTGAGACGATGGGGTC 103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
218 UMetPheArgProAspProAlaMetLeuSerLeuGluAsnLysTrpSerA 235
104 GACAAATCCACCAAGTCACAGATCC...AAGATGSCATTTTAAATAC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
235 lAluSerAsnAlaValGlnGlnSerAsnLysProValTyrValGlySerVa 268
151 CAATTAAGATTGCAATTCGCCGCGTAACTTTGTGACGNTGACAAATG 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
252 ProValGluLeuArgSerIleSerAsnLysProValTyrValGlySerVa 268
201 TCGCGAC...GTGATCTCCAGCTTGGGATGATGTTGTTGATGATGC... 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 lSerAspArgValIleSerGlyLeuAlaIleMetLeuPheIleCysArgS 285
248 GTGGGCAATTCCTCT.....CTCGACACCCCTTCGCGCGT 285
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
285 erThAspArgAlaSerSerAspGlnPheIleAspHis.....LeuLeu 299
286 CTCCTAAGTCCGTCGTCG.....GATCGCGC 311

```

```

seq_name: sp_plant:094BW5
seq. documentation block:
ID Q94BW5 PRELIMINARY; PRT; 581 AA.
AC Q94BW5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TYPE 2 RIBOSOME-INACTIVATING PROTEIN CINNAMOMIN I PRECURSOR.
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomin proteins and study of their expression
RT patterns."
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AY039801; AAK62458.1; -.
DR SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 581 TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOMIN I.
FT
FT SEQUENCE 581 AA; 64215 MW; 6E8F5FB8BA3D196 CRC64;

```

```

alignment_scores:
  Quality: 179.00 Length: 131
  Ratio: 2.131 Gaps: 5
Percent Similarity: 64.122 Percent Identity: 38.931

```

```

alignment_block:
US-09-627-165B-13 x Q94BW5 ..

```

Align seg 1/1 to: Q94BW5 from: 1 to: 581

```

4 AGATTCAATCCCATCCTGTCGAGGCTTCCGCGCAATTAACAGTGGGGA 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
202 ArgPheArgPheIleGluTyrArgValArgLysSerIleSerArgAlaG 218
54 GTCNCTCTCCCAACATGATGATGATGATGATGATGATGATGATGATG 103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
218 UmetPheArgProAspProAlaMetLeuSerLeuGluLysSerIlePhe 235
104 GACAAATCCCAACAGTCCAGCAGTCC...AAGATGCAATTTTAATAC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
235 IaleuSerAsnAlaValGlnGlnSerAsnGlnGlyAlaPheSerSer 251
151 CAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
252 ProValGluLeuArgSerIleSerAsnLysProValTyrValGlySer 268
201 TCCGCGAC...GTGATCTCAGCTTGGGATGATGATGATGATGATGATG 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 IserAspArgValIleSerGlyLeuAlaIleMetLeuPheIleCysArg 285
248 GTCGCGCATTCCTCTCT...CTGACCAACCTTCGCGCGTC 285
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
285 erThrAspArgAlaSerSerAspGlnPheIleAspHis...MetLeu 299
286 CTCCTAAGTCCGTCGTG...GATCGCGC 311
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
300 MetIleArgProIleLeuValAspValAlaGluValAlaThrAspAla 316

```

```

seq_name: sp_plant:Q41174
seq. documentation block:
ID Q41174 PRELIMINARY; PRT; 541 AA.
AC Q41174;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RRNA - GLYCOSIDASE (EC 3.2.2.22) (FRAGMENT).
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-92338377; PubMed-163311;
RX Roberts L.M., Tregear J.W., Lord J.M.;
RT "Molecular cloning of ricin."
RT Targeted Diagn. Ther. 7:81-97(1992).
CC -1- CATALYTIC ACTIVITY: ENDOPHYTOXINS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
DR EMBL: S40366; AAB22582.1; -.
DR HSRP: P02879; IBR6.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR InterPro: IPR001400; SOMATOTROPIN.
DR Pfam: PF00652; Ricin_B_lectin; 2.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; UNKNOWN_1.
KW Hydrolyase; Toxin.
FT NON_TER 1 1
FT SEQUENCE 541 AA; 60281 MW; 2B7B2CDEF1F2E9D9 CRC64;

```

```

alignment_scores:
  Quality: 178.00 Length: 119
  Ratio: 2.342 Gaps: 1
Percent Similarity: 63.866 Percent Identity: 36.975

```

```

alignment_block:
US-09-627-165B-13 x Q41174 ..

```

Align seg 1/1 to: Q41174 from: 1 to: 541

```

1 GCCAGATTCATCCCATCCTGTCGAGGCTTCCGCGCAATTAACAGTGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179 AlaArgPheArgPheIleGluTyrArgValArgLysSerIleSerArg 195
51 GGAGTCTCTCCCAACATGATGATGATGATGATGATGATGATGATGAT 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
195 nArgArgSerAlaProAspProSerValIlePheLeuGluLysSerTyr 212
101 GTCGCAATTCACCAACAGTCCAGCAGTCCAGATGATGATGATGATGAT 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
212 LyrArgLeuSerThrAlaIleGlnGlnSerAsnGlnGlyAlaPheSer 228
151 CAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
229 ProIleGluLeuGlnArgArgAsnGlySerLysPheSerValTyrAsp 245
201 TCCGCGATGATCTCCAGCTTGGGATGATGATGATGATGATGATGATG 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
245 IserIleLeuPheProIleLeuValAspValAlaGluValAlaThrAsp 262

```

251 GGCCATCTCTCTCTGACACCCCTTGGCGCTGCTCTAAGTCGTC 300  
 ||| |||:||||| |||:||||| |||  
 262 rprfserferlnpheser.....leuilelgrproval 274  
 ||| |||:||||| |||:||||| |||  
 301 GTGATGCGGCCACGATGTCACCTGACACTNTTCCGACCCACCGTCG 350  
 ||| |||:||||| |||:||||| |||  
 275 ValProaenPhenaAlaAspValcysMetAspProgluProlleValar 291  
 ||| |||:||||| |||:||||| |||  
 351 CATCGTA 357  
 ||| |||:||||| |||:||||| |||  
 291 glileval 293

seq\_name: sp\_plant:Q06076

seq\_documentation\_block:  
 ID Q06076 PRELIMINARY; PRT; 528 AA.  
 AC Q06076;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE RNA - GLYCOSIDASE (EC 3.2.2.22) (FRAGMENT)  
 OS Atrius precatorius (Indian licorice) (Crab's eye)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Adreae; Atrius.  
 ON NCBI\_TaxID=3816;  
 RN |||  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93132798; PubMed=8421313;  
 RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;  
 RT "Primary structure of three distinct isoabins determined by cDNA  
 sequencing: conservation and significance";  
 RL J. Mol. Biol. 229:263-267(1993).  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.  
 EMBL: M98346; AAA32626.1; -.  
 DR HSRP; P11440; IABR.  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 2.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; UNKNOWN\_1.  
 KW Hydrolase; Toxin.  
 FT NON\_TER 1 528  
 FT 528  
 SQ SEQUENCE 528 AA; 58870 MW; 62BD42FB8FBE60F8 CRC64;

alignment\_scores:

Quality: 139.00 Length: 119  
 Ratio: 1.759 Gaps: 4  
 Percent Similarity: 66.387 Percent Identity: 36.134

alignment\_block:

US-09-627-165b-13 x Q06076 ..

Align seg 1/1 to: Q06076 from: 1 to: 528

1 GCCAGATTCATCCCATGTCGAGCGCTGCGCGCAATTACAGTGG 50  
 |||:||||| |||:||||| |||:||||| |||  
 166 AlArGtYrArGcYsIlSeSerAsnArGValGlyValSerIlArGThGl 182  
 ||| |||:||||| |||:||||| |||:||||| |||  
 51 GGAGTNTCTCCACCAACATGATGATCGAGTGCAGTGCAGTGG 100  
 ||| |||:||||| |||:||||| |||:||||| |||  
 182 yTrAlaPhaPheGlnProAspProAlaMetLeuSerLeuGluAsnArGTrpa 199  
 ||| |||:||||| |||:||||| |||:||||| |||  
 101 GTGGAATTCACCAAGTCAGAGATCC...AAGATGGCATTTTAAAT 147  
 ||| |||:||||| |||:||||| |||:||||| |||

199 spAsnLeuSerGlyValAlaGlnInSerValAlaAspAlaPheProAsn 215  
 148 ACCCAATAGATTGACGATTTCCGCCGTAACCTTGTGACGNTGACCA 197  
 ||| |||:||||| |||:||||| |||:||||| |||  
 216 AsnValIlleuSerSerIlleAsnArGlnProValValaValAspSorLe 232  
 ||| |||:||||| |||:||||| |||:||||| |||  
 198 TGTTCGACGATGATCTCCACCTTGGCGAWACATGTTGTTCAGATGAGT 247  
 ||| |||:||||| |||:||||| |||:||||| |||  
 232 uSerIlSProThrValAlaValleuAlaLeuMetLeuPheValcysAsn. 248  
 ||| |||:||||| |||:||||| |||:||||| |||  
 248 GTGCGCATCTCTCTCTGACACCCCTTGGCGCTGCTCTAAGTCG 297  
 ||| |||:||||| |||:||||| |||:||||| |||  
 249 ....ProProAsnAla.....AsnInSerProleuIleArGSer 261  
 ||| |||:||||| |||:||||| |||:||||| |||  
 298 GTGATGATGCGGCCACGATGTCACCTGACACTNTTCCGACCCACCGT 347  
 ||| |||:||||| |||:||||| |||:||||| |||  
 262 lIeValGlGlnSerLys...lIeCysSerSerArGTrGluProThrVa 277  
 ||| |||:||||| |||:||||| |||:||||| |||  
 348 GCGCATC 354  
 ||| |||:||||| |||:||||| |||:||||| |||  
 277 lArGtle 279

seq\_name: sp\_plant:Q9M653

seq\_documentation\_block:  
 ID Q9M653 PRELIMINARY; PRT; 603 AA.  
 AC Q9M653;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE RNA - GLYCOSIDASE (EC 3.2.2.22).  
 GN R1PT.  
 OS Polygonatum multiflorum.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;  
 OC Convallariaceae; Polygonatum.  
 ON NCBI\_TaxID=45371;  
 RN |||  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20247177; PubMed=10785398;  
 RA Van Damme E.J., Hao Q., Charrels D., Barre A., Rouge P., Van Leuven F.,  
 RA Peumans W.J.;  
 RT "Characterization and molecular cloning of two different type 2  
 RT ribosome-inactivating proteins from the monocotyledonous plant  
 RT Polygonatum multiflorum";  
 RL Eur. J. Biochem. 267:2746-2759(2000).  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.  
 EMBL: AF213984; AF37219.1; -.  
 DR HSRP; P02879; ZAAI.  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; UNKNOWN\_1.  
 KW Hydrolase; Toxin.  
 SQ SEQUENCE 603 AA; 66733 MW; 88B61C17CA431B1E CRC64;

alignment\_scores:

Quality: 128.50 Length: 140  
 Ratio: 1.606 Gaps: 7  
 Percent Similarity: 57.143 Percent Identity: 31.429

alignment\_block:

US-09-627-165b-13 x Q9M653 ..

Align seg 1/1 to: Q9M653 from: 1 to: 603

```

1 GCCAGATTCAATCCATCCTGTCGAGCGTCCCGGCAATTAAACAGTGG 50
  |||||
210 AlaArgPheArgHisIleGluIArgValArgArgSerPheThrAlaAl 226
51 GGAGTCNTCTCCA.....CCAAACATGTACATGCTCGAGCTGG 88
  |||||
226 AASPHisPheIleuThrPheArgProAspGlyLeuMetLeuSerMetG 243
89 AGACGAGTGGGTCGACAAATCCCAAGTCAGCAGTCC...AAGGAT 135
  |||||
243 LuAsnGluTrpProSerMetSerLeuGluValGlnArgSerIleGluGly 259
136 GGCATTTTATACCAATAGATTCAG.....AT 167
  |||||
260 GlyIlePheIleGlyValIleGlnLeuGlnAspIleSerPheGlnPro 276
168 TTCGCGCGTAACTTGTGACAGNTGACGATGTCGCGAGCTGATCCGA 217
  |||||
276 ValArgValAspAsnPheAsnThrLeuSerArg.....TyrThrM 289
218 GCTTGGCGCATGTTGTCGATGACGAGTGGCGCA.....AT 255
  |||||
289 etValAlaLeuLeuLeuPheArgCys...GlyHisProArgAlaThrAla 304
256 ...TTCCTCTCTCGACACCCCTCGCGCTGCTCCTAAGTCCGCT 302
  |||||
305 GlyThrSerSerThrThrProAlaIleAlaGlnIleIleArgMetProva 321
303 GGATGGCGCC.....AAGCATGTCACCTGACCTGACTTTT 334
  |||||
321 AspValIleuAlaGlyGluGluIleTyrThrAspGluGluThrCysThrValG 338
335 CCGAACCCACCGTGGCGCATC 354
338 LysGluProThrArgArgIle 344

```

seq\_name: sp\_plant:Q9M654

seq\_documentation\_block:

ID Q9M654 PRELIMINARY; PRT: 604 AA.

```

AC Q9M654;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RNA - GLYCOSIDASE (EC 3.2.2.22).
GN R1PM.
OS Polygonatum multiflorum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
OC Convallariaceae; Polygonatum.
OC NCBI_TaxID=45371;
OX [1]
RN RP
RX MEDLINE-20247177; PubMed-10785398;
RA Van Damme E.J., Hao Q., Charrels D., Barre A., Rouge P., Van Leuven F.,
Peumans W.J.;
RT "Characterization and molecular cloning of two different type 2
RT ribosome-inactivating proteins from the monocotyledonous plant
RT Polygonatum multiflorum".
RL Eur. J. Biochem. 267:2746-2759(2000).
CC -1- CATALYTIC ACTIVITY: ENDORHYDOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
EMBL: AF213983; AAF37218.1; -.
DR HSSP; P02879; 2AAT.
DR InterPro; IPR00772; R1cin_B_lectin.
DR InterPro; IPR001574; R1P.
DR Pfam; PF00652; R1cin_B_lectin; 5.
DR Pfam; PF00161; R1P; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; UNKNOWN_1.

```

KW Hydrolyase; Toxin.  
SQ SEQUENCE 604 AA; 67060 MW; 6206C2B25F90B4E CRC64;

alignment\_scores: Length: 136  
Quality: 128.50 Gaps: 6  
Ratio: 1.669  
Percent Similarity: 56.618 Percent Identity: 30.882

alignment\_block:  
US-09-627-165B-13 x Q9M654 ..

Align seg 1/1 to: Q9M654 from: 1 to: 604

```

1 GCCAGATTCAATCCATCCTGTCGAGCGTCCCGGCAATTAAACAGTGG 50
  |||||
213 AlaArgPheArgHisIleGluIArgValArgArgSerIleAlaAspG 229
51 GGAGTCNTCTCCACCAACATGTACATGCTCGAGCTGAGCAGGTTGG 100
  |||||
229 nAspThrPheThrProGlySerLeuMetIleThrMetGluValysTrp 246
101 GTCGCAATCCACCAAGTCCAGCAGTCCAAGAT..... 135
  |||||
136 ...GGCATTTTAAATACCAATAGATTG.....CA 164
246 etLysMetSerGlnGlnValGluArgSerValAsnAspGlnGlyIlePhe 262
263 ThrGlyIlePheThrArgThrValGlnLeuIleAspAsnLeuGlnTh 279
165 GATTTCCGCGCGTAATGTTGTGACGNTGACGATGTCGCGCAGTGA 214
  |||||
279 rLeuAsnIleAspAsnPe.....AsnAlaLeuSerLeuHisT 292
215 CCAAGTGGCGCATGTTGTTGCAATGCAATGTCGCGCATTCCTCT 264
  |||||
292 hMetLeuAlaIleLeuLeuPheArgCys.....ArgThrThrArgSer 306
265 CTCGACCAACCCCTCGCGG.....CTGCTCTAAGTCT 296
  |||||
307 SerHisAsnThrLeuProAlaIleSerAsnIleValLeuMetGlyGluAs 323
297 CGTGTGAGATGCGCGCAAGCATGTACCTGCACTGTTTCGGAACCCAG 346
  |||||
323 pTyrValAsp...LysAspAspGluIysCysThrValGlyGluProTha 339
347 TGGCGCAT 354
  |||||
339 rGArgIle 341

```

seq\_name: sp\_plant:Q9M6E9

seq\_documentation\_block:

ID Q9M6E9 PRELIMINARY; PRT: 547 AA.

```

AC Q9M6E9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RNA - GLYCOSIDASE (EC 3.2.2.22).
GN AAG.
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Adreace; Abrus.
OC NCBI_TaxID=3816;
OX [1]
RN RP
RX MEDLINE-20102702; PubMed-10636890;
RA Liu C.L., Tsai C.C., Lin S.C., Wang L.I., Hsu C.I., Hwang M.J.,
Lin J.Y.;
RT "Primary Structure and Function Analysis of the Abrus precatorius
RT Agglutinin A Chain by Site-directed Mutagenesis: Pro199 of Amphiphilic
RT alpha-Helix H Impairs Protein Synthesis Inhibitory Activity."

```





```

seq_documentation_block:
ID 022415 PRELIMINARY: PRT: 570 AA.
AC 022415:
DT 01-JAN-1998 (TRENBLREL. 05, Created)
DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE RNA - GLYCOSIDASE PRECURSOR (EC 3.2.2.22).
GN LEC5NA-1F.
OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Adoxaceae; Sambucus.
OC NCBI_TaxID=4202;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FRUIT;
RC MEDLINE=98200071; PubMed=9541002;
RA Peumans W.J., Roy S., Baire A., Rouge P., van Leuven F.,
RA van Damme E.J.M.;
RT "Elderberry (Sambucus nigra) contains truncated Neu5Ac(alpha-
RT 2,6)Gal/GalNAc-binding type 2 ribosome-inactivating proteins.";
RL FEBS Lett. 425:35-39(1998).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
CC EMBL: AF012899; AAC49989.1; -.
CC HSSP: P02879; ZNAI.
DR InterPro: IPR000772; R1cin_B_lectin.
DR InterPro: IPR001574; R1P.
DR Pfam: PF00652; R1cin_B_lectin; 2.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; R1CIN.
DR PROSITE: PS50231; R1CIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_R1CIN; 1.
KW Hydrolyase; Signal; Toxin.
FT SIGNAL 1 28
FT CHAIN 29 308
FT CHAIN 309 570
FT CHAIN TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CHAIN SNAIF, B CHAIN.
SQ SEQUENCE 570 AA; 62752 MW; 7D799D3A20764BA6 CRC64;

alignment_scores:
Quality: 119.50 Length: 127
Ratio: 1.475 Gaps: 4
Percent Similarity: 63.780 Percent Identity: 25.197

alignment_block:
US-09-627-165B-13 x 022415 ..
Align seg 1/1 to: 022415 from: 1 to: 570
1 GCCAGATTCATCCCATCCTGAGGCTTGGCCGCAATTAACAGTGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 Alaargpneargyrllegluleuarglleargthserlethrspal 217
51 GGAGTCNTCTCCACCAACATGTCATGTCGAGTCGACGAGTGGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
217 aserlupherthrproasplaleuauetleusermetclunasntrps 234
101 GTGCACATTCACCCAGTCAGCAGTCCAGG...GATGSCATTTTAA 147
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
234 erSerMetSerSerSerlulleglnlalnaglnproglyglylephpro 250
148 ACCCAATATAGATTCAGATTTCCCGCGTAACCTTGTGACGNTGACAA 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 glyvalvalgluleuargspglulrgnshsnproilleglualthrAs 267
198 TGTTCCGACAGTG.....ATCTCCAGCTTGCGCATGTGTTGCAT 241
| ||| :||| :|||:|||||:|||||:|||||:

```

```

267 nPheargargleuphegluleuThrTyrillealavalLeuLeuTyrGlyC 284
242 GCAGTGTGGCCATCTCTCTGACACCCCTTGCCTGCTCCTA 291
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
284 yAla.....ProValThrSerAsnSerTyrThrAsnAlaIleasp 298
292 AGTCGCTGCTG.....GATCGCGCCACAGATGAC 323
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 AlaglnlelleuysmetProValPheargglyglyTyrGlyulayva 315
324 CTGCATCTTTCCGACCCACCTGCGCATC 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 lCySserValvalglulvalthrargArgyle 325

seq_name: sp.Plant:004367
seq_documentation_block:
ID 004367 PRELIMINARY: PRT: 563 AA.
AC 004367:
DT 01-JUL-1997 (TRENBLREL. 04, Created)
DT 01-JUL-1997 (TRENBLREL. 04, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE RNA - GLYCOSIDASE PRECURSOR (EC 3.2.2.22).
GN Sambucus nigra (European elder).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Adoxaceae; Sambucus.
OC NCBI_TaxID=4202;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98112023; PubMed=9450339;
RA Van Damme E.J., Roy S., Baire A., Rouge P., Van Leuven F.,
RA Peumans W.J.;
RT "The major elderberry (Sambucus nigra) fruit protein is a lectin
RT derived from a truncated type 2 ribosome-inactivating protein.";
RL Plant J. 12:1251-1260(1997).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
CC EMBL: U76524; AAC15886.1; -.
CC HSSP: P02879; ZNAI.
DR InterPro: IPR000772; R1cin_B_lectin.
DR InterPro: IPR001574; R1P.
DR Pfam: PF00652; R1cin_B_lectin; 2.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; R1CIN.
DR PROSITE: PS50231; R1CIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_R1CIN; 1.
KW Hydrolyase; Signal; Toxin.
FT SIGNAL 1 25
FT CHAIN 26 297
FT CHAIN 298 563
FT CHAIN RIBOSOME INACTIVATING PROTEIN, A CHAIN.
SQ SEQUENCE 563 AA; 62336 MW; 3ED2B6C0B796205 CRC64;

alignment_scores:
Quality: 109.00 Length: 89
Ratio: 1.879 Gaps: 3
Percent Similarity: 65.169 Percent Identity: 32.584

alignment_block:
US-09-627-165B-13 x 004367 ..
Align seg 1/1 to: 004367 from: 1 to: 563
1 GCCAGATTCATCCCATCCTGAGGCTTGGCCGCAATTAACAGTGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
190 Alaargpneargyrlleglulnaglnlalnaglnproglyglylephpro 206
51 GGAGTCNTCTCCACCAACATGTCATGTCGAGTCGACGAGTGGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 ralaglypnerthrproasplalaleuauetleusermetclunasntrps 223

```











PF 27-JUL-2000; 2000EP-0402168.  
 XX 27-JUL-1999; 99KR-0030638.  
 PR (MIST-) MISTLE BIOTECH CO LTD.  
 XX Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;  
 PI Kang T, Park C;  
 XX MPI; 2001-171044/18.  
 DR N-PSDB; AAC85472.  
 XX  
 PT Novel lectin proteins isolated from Korean mistletoe, useful for  
 XX enhancing immunity and effectuating anti-tumoral activity -  
 PS Claim 2; Page 25-26; 62pp; English.  
 XX  
 CC The sequences given in ABA47090-92 are isoforms of an A-chain gene  
 CC involved in biosynthesis of lectins isolated from Korean mistletoe.  
 CC Korean mistletoe lectins (KML) are useful for enhancing immunity and  
 CC for treating tumours. The KML's are isolated from a protein fraction  
 CC derived from the leaves, stems and fruits of Korean mistletoe, which  
 CC is designated KM-110. One of the isolates, KML-C was shown to be  
 CC effective against colon 26-M3.1 carcinoma and L5178Y-M25 lymphoma.  
 XX  
 SO Sequence 254 AA;

alignment\_scores: Length: 140  
 Quality: 620.00 Gaps: 0  
 Ratio: 4.593  
 Percent Similarity: 96.429 Percent Identity: 87.857

alignment\_block:  
 US-09-627-165B-15 x ABA47090 ..

Align seg 1/1 to: ABA47090 from: 1 to: 254

```

1 TACACAGATCTGAGCGATACGGCGTCATAGGACGAGATCCCTGGG 50
  |||||||
115 TGTThrsprleugluarphnealaglyhlsarqspelnlleproleugl 131
  |||||||
51 TATAGGAACTCATTCATCCGTCTCGGGCTTATTCAGGCGGCA 100
  |||||||
131 YARGluileulleeginservalserAlaleuArghpProglYsera 148
  |||||||
101 GCACCCGGGCCCAAGCTGCTTCCTTATATCCATTCAGATGATCC 150
  |||||||
148 sntThrArgAlaglInAlarGserPhelIetIleuIleGImeIleSer 164
  |||||||
151 GAGGCGCGGAGATTCATCCATCTGTCGGGCTTATTCAGGCGCA 200
  |||||||
165 GluAlAlAlArghpneAsnProIleuItrpArghAlArGInYrIleSe 181
  |||||||
201 CAGCGGGGAGTATTTCTCCCGCATGTACATGCTGAGCTGAGACTA 250
  |||||||
181 rserglYgIserPhelIeuProAspThrYrIleuGInleuGInuTrhS 198
  |||||||
251 GTTGGGGCCCAACATCCAGCAAGTCCAGAGCTTACGATCGCGTTT 300
  |||||||
198 eTrIpelYgInGInserThrGInvalGInhIserThAspIlyAlpHe 214
  |||||||
301 AATTAACCATTTGGTGGGTATATGCACCGGTAACCTTCGTAAGT 350
  |||||||
215 AsnAsnProIleArghleuThrIleSerThrGlyValPheValThrIleuSe 231
  |||||||
351 CAATGTCGGAGAGTGATGCGCAGCTTAGCATGATGTTGTATGTA 400
  |||||||
231 rAsnValArghspValIleAlaSerIleuAlaIleMetIleuPheValCyG 248
  |||||||
401 GGGACGACCATTTCTCTCC 420
  |||||||
248 lAspArghProserSerSer 254
  |||||||

```

seqname: /STDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:AAW64661

seq\_documentation\_block:  
 ID AAW64661 standard; Protein: 252 AA.

XX AAW64661;

XX 23-OCT-1998 (first entry)

XX Mistletoe rMLA variant protein.

XX Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;  
 XX intracellular; processing module; protease recognition; targeting module;  
 XX internalisation; treatment; disorder; cell proliferation; activation;  
 XX autoimmune disease; allergy; tumour; ricin; translocation; ss.

XX Viscum album.

XX Key Location/Qualifiers  
 XX Protein 1..252  
 XX /note="partial"

XX W09829540-A2.

XX 09-JUL-1998.

XX 02-JAN-1998; 98WO-EP00009.

XX 02-JAN-1997; 97EP-0100012.

XX (BRA1-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.

XX Eck J, Schmidt A, Zinke H;

XX MPI; 1998-388122/33.

XX N-PSDB; AAW51343.

XX Nucleic acid encoding fusion protein containing mistletoe lectin A  
 XX chain - useful for treatment of proliferative and autoimmune  
 XX diseases, allergies and tumours  
 XX  
 XX Disclosure: Flg 11a'; 115pp; German.

XX This sequence encodes a variant mistletoe lectin A-chain, rMLA. This  
 XX sequence can be used in the construction of a fusion protein which  
 XX comprises an effector module that is cytotoxic intracellularly, a  
 XX processing module covalently bonded to the effector module and  
 XX containing a protease recognition sequence, and a targeting module  
 XX covalently bonded to the processing module, able to bind specifically to  
 XX the surface of a cell so as to mediate internalisation of the fusion  
 XX protein. Such a fusion protein can be used for treating disorders  
 XX involving proliferation and/or elevated activation of cells, especially  
 XX autoimmune disease, allergy and tumours. The proteins can be administered  
 XX e.g. by injection or topically but especially by intravenous injection,  
 XX at 1 ng to 500 mu g/kg/day, or for ex vivo use at 1 pg to 500 ng/ml.  
 XX Fusion proteins can develop toxic activity in a wide range of target  
 XX cells. The processing module prevents extracellular dissociation, and  
 XX fusion proteins based on mistletoe lectin A-chain are far more active  
 XX than those based on ricin and do have the associated problems of  
 XX non-specific toxicity. The protein may be expressed in a non-glycosylated  
 XX form that does not bind to sugar receptors in the liver, and which has a  
 XX long half-life in the blood. Where the mistletoe lectin B-chain is used,  
 XX it actively assists in translocation of the ML A-chain from the  
 XX endoplasmic reticulum to the cytoplasm.

SO Sequence 252 AA;

alignment\_scores: Length: 138  
 Quality: 619.00 Gaps: 0  
 Ratio: 4.654  
 Percent Similarity: 96.377 Percent Identity: 86.232

alignment\_block:  
US-09-627-165b-15 x AAW64661 ..

Align seg 1/1 to: AAW64661 from: 1 to: 252

```

1  TACACAGATCTGTGAGCATACCGCGGTCAATAGGAGGACAGATCCCTGTGG 50
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115  TYPProaspLeucluarGTYrAlaGlyHISArGAspGlnIleProLeuG1 131
51  TATAGAGAACTCATTCGATCCGCTCGCGGCTTCGTTATCCAGCGGCA 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131  YILeaspGlnIleuIleGlnSerValIThrAlaLeuAArgProGlyGlys 148
101  GCACCGGGGCCCAAGCTCGTCCCTTAATTCCTCATTCAGATCTCC 150
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148  eRThrArgThrGlnAlaArgSerIleLeuIleLeuIleGlnMetIleSer 164
151  GAGCGCGGAGATCATCCCATCTTTTGGAGGGCTCGCAATACATTAA 200
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165  GLuAlaAlaArgPheAsnProIleLeuThrArgAlaArgGlnThrIleAs 181
201  CACCGGGGAGTCATTTCTCCGACATGTACATGCTCGAGCTCGAGACTA 250
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181  nserGlyAlaSerPheLeuProAspValIYrMetLeuGlnLeuGlnTrs 198
251  GTTGGGGGCCCAACATCCACGCAAGTCTCAGAGTCTCAGGATGCGTTT 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198  eRTrpGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyAlaPhe 214
301  AATAACCATTTGCGTGGTATATCCACCGGTACTCTGTCGAGCTGAG 350
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215  AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValIThrLeuH 231
351  CAATGTCGCGAGTGCATCCGAGTACGATCATGCTGTTGTATGTA 400
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
231  rAsnValArGAspValIleAlaSerLeuAlaIleMetLeuPheValCysG 248
401  GGGACCGACCATCT 414
   ::||| |||||
248  IYglnArgProSer 252

```

seq\_name: /stid1/gcgdata/hold--geneseq/geneseq-emb1/AA1997.DAT:AAW10022

seq\_documentation\_block:

ID AAW10022 standard; Protein; 253 AA.

XX AAW10022;

DT 18-DEC-1997 (first entry)

DE Prepro mistletoe lectin A chain.

KW Mistletoe; lectin; cytotoxic; A chain; B chain; dimer.

OS Viscum album.

PN RP751221-A1.

PD 02-JAN-1997.

XX 26-JUN-1995; 95EP-0109949.

PR 26-JUN-1995; 95EP-0109949.

XX (MADU ) MADUUS KOELN AG.

PI Baur A, Eck J, Lentzen H, Zinke H;

DR WPI: 1997-054678/06.

DR N-PSDB; AAT70474.

PT Nucleic acid encoding pre-pro form of mistletoe lectin - for

PT therapeutic or diagnostic use

XX Claim 12; Fig 4A; 30pp; German.

CC Mistletoe lectin is a cytotoxic agent that has been used for tumour

CC therapy. It can be used in immunotoxins and medicaments. Nucleic

CC acid fragments can be used in diagnostic methods. Mistletoe lectin (

XX AAT70473) comprises an A chain (AAT70474) and a B chain (AAT70475).

XX Sequence 253 AA:

alignment\_scores:  
Quality: 619.00 Length: 138  
Ratio: 4.654 Gaps: 0  
Percent Similarity: 96.377 Percent Identity: 86.232

alignment\_block:  
US-09-627-165b-15 x AAW10022 ..

Align seg 1/1 to: AAW10022 from: 1 to: 253

```

1  TACACAGATCTGTGAGCATACCGCGGTCAATAGGAGGACAGATCCCTGTGG 50
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116  TYPProaspLeucluarGTYrAlaGlyHISArGAspGlnIleProLeuG1 132
51  TATAGAGAACTCATTCGATCCGCTCGCGGCTTCGTTATCCAGCGGCA 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132  YILeaspGlnIleuIleGlnSerValIThrAlaLeuAArgProGlyGlys 149
101  GCACCGGGGCCCAAGCTCGTCCCTTAATTCCTCATTCAGATCTCTCC 150
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149  eRThrArgThrGlnAlaArgSerIleLeuIleLeuIleGlnMetIleSer 165
151  GAGCGCGGAGATCATCCCATCTTTTGGAGGGCTCGCAATACATTAA 200
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
166  GLuAlaAlaArgPheAsnProIleLeuThrArgAlaArgGlnThrIleAs 182
201  CACCGGGGAGTCATTTCTCCGACATGTACATGCTCAGAGTCTGAGACTA 250
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182  nserGlyAlaSerPheLeuProAspValIYrMetLeuGlnLeuGlnTrs 199
251  GTTGGGGGCCCAACATCCACGCAAGTCTCAGAGTCTCAGGATGCGTTT 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199  eRTrpGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyAlaPhe 215
301  AATAACCATTTGCGTGGTATATCCACCGGTACTCTGTCGAGCTGAG 350
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216  AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValIThrLeuH 232
351  CAATGTCGCGAGTGCATCCGAGTACGATCATGCTGTTGTATGTA 400
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232  rAsnValArGAspValIleAlaSerLeuAlaIleMetLeuPheValCysG 249
401  GGGACCGACCATCT 414
   ::||| |||||
249  IYglnArgProSer 253

```

seq\_name: /stid1/gcgdata/hold--geneseq/geneseq-emb1/AA1999.DAT:AAW90125

seq\_documentation\_block:

ID AAW90125 standard; Protein; 253 AA.

XX AAW90125;

DT 30-APR-1999 (first entry)

DE Mistletoe ML A-chain protein.

KW ML; mistletoe; lectin; ML; A-chain; transgenic plant; glycosylation;

KW dimer; immunotoxin; large-scale production; diagnosis; therapeutic;

XX cancer.







```
seq_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT: AAB47092
seq_documentation_block:
ID   AAB47092 standard; Protein; 256 AA.
XX
AC   AAB47092;
XX
DT   16-MAY-2001 (first entry)
XX
DE   A-chain isoform for biosynthesis of a Korean mistletoe lectin #3.
XX
KW   Isoform; A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
KW   KML; tumour; KM-110; KML-C; KMHP; KML-IIU; KML-III;
KW   heparin binding protein.
XX
OS   Viscum album coloratum.
XX
PN   EPI074560-A2.
XX
PD   07-FEB-2001.
XX
PE   27-JUL-2000; 2000EP-0402168.
XX
PR   27-JUL-1999; 99KR-0030638.
XX
PA   (MIST-) MISTLE BIOTECH CO LTD.
XX
PI   Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
PI   Kang T, Park C;
XX
DR   WPI: 2001-171044/18.
XX
DR   N-PSDB: AAC85474.
XX
PT   Novel lectin proteins isolated from Korean mistletoe, useful for
PT   enhancing immunity and effectuating anti-tumoral activity -
XX
PS   Claim 6; Page 28; 62pp; English.
XX
CC   The sequences given in AAB47090-92 are isoforms of an A-chain gene
CC   involved in biosynthesis of lectins isolated from Korean mistletoe.
CC   Korean mistletoe lectins (KML) are useful for enhancing immunity and
CC   for treating tumours. The KML's are isolated from a protein fraction
CC   derived from the leaves, stems and fruits of Korean mistletoe, which
CC   is designated KM-110. One of the isolates, KML-C was shown to be
CC   effective against colon 26-M3.1 carcinoma and L5178Y-M25 lymphoma.
XX
SQ   Sequence 256 AA;

alignment_scores:
      Quality: 611.00      Length: 139
      Ratio: 4.629      Gaps: 0
      Percent Similarity: 94.964      Percent Identity: 86.331

alignment_block:
US-09-627-165B-15 x AAB47092 ..

Align seg 1/1 to: AAB47092 from: 1 to: 256

1 TACACAGATCTGAGCGATAGCGCGGTATAGGACCAAGATCCCTCTGGG 50
|||||
117 TYRALAASPLeuGluArgTyrAlaGlyHisArgAspArgIleProLeuG1 133
51 TATNAGGAGAACTCATTCATCTCGCTCGCGCTTCGTTATCCAGCGGCA 100
|||||
133 yArgGluProLeuIleArgSerValSerAlaLeuAspTyrProGlyG1S 150
101 GCACCCGGGCGCAAGCTCGTCCCTTATATCTCTATTCAGATGATCC 150
|||||
150 erThrArgAlaGlnAlaSerSerIleIleIleValIleIleGlnMetIleSer 166
151 GAGCGCGCGAGTTCATCCATCTTTGGAGGGCTCGGCATACATTAA 200
```

```
|||||
167 GluAlaIleArgPheAsnProIleLeuTrpArgAlaArgGlnTyrIleAs 183
201 CAGCGGGAGCTCATTTCTTCCGACATGTACATGCTCGAGCTGAGACTA 250
|||||
183 nSerGlyValSerTyrLeuProAspValTyrMetLeuGluLeuGluAlaS 200
251 GTTGGGGCACAACATCCAGCAGATCCAGCAGCTCTACGATGGCGTTT 300
|||||
200 erTrpGlyGlnInserThrGlnValGlnInserThrAspGlyValPhe 216
301 AATAACCATTTGGTTGGTATATCCACCGGTAACTTGCTGACGTTGAG 350
|||||
217 AsnAsnProIleArgLeuGlyIleSerThrGlyAsnPheValTrpLeuS 233
351 CAATGTCGAGCGAGTACGCGACCTTAGCGATCATGTTGTTGATGTA 400
|||||
233 rAsnValArgAspValIleAlaSerIleuGlyIleMetValPheValCysa 250
401 GGGACCGACATCTTCC 417
|||||
250 rgAspArgSerSerSer 255

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT: AA25971
seq_documentation_block:
ID   AA25971 standard; Protein; 255 AA.
XX
AC   AA25971;
XX
DT   18-OCT-1999 (first entry)
XX
DE   Mistletoe lectin A chain MLA consensus protein sequence 1.
XX
KW   Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
KW   ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW   lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW   cancer; cytotoxicity; antigen; isoform.
XX
OS   Viscum album.
XX
FH   Key Location/Qualifiers
FT   Misc-difference 15 /label= Asp, Glu
FT   FT
FT   Misc-difference 63 /label= Gly, Gln
FT   FT
FT   Misc-difference 66 /label= Ile, Val
FT   FT
FT   Misc-difference 75 /label= Leu, Ala
FT   FT
FT   Misc-difference 107 /label= Asp, Arg or none
FT   FT
FT   Misc-difference 113 /label= Asn, Thr
FT   FT
FT   Misc-difference 117 /label= Pro, Thr
FT   FT
FT   Misc-difference 134 /label= Asp, Glu
FT   FT
FT   Misc-difference 141 /label= Ser, Thr
FT   FT
FT   Misc-difference 145 /label= Phe, Tyr
FT   FT
FT   Misc-difference 152 /label= Thr, Ala
FT   FT
FT   Misc-difference 177 /label= Ala, Tyr
FT   FT
FT   Misc-difference 180 /label= Tyr, Asp
FT   FT
FT   Misc-difference 185 /label= Ala, Glu
FT   FT
FT   Misc-difference 191 /label= Val, Met
FT   FT
FT   Misc-difference 219
```

```

FT      /label= Ile, Phe
FT      Misc-difference 224
FT      /label= Pro, Ser
FT      Misc-difference 225
FT      /label= Pro, Thr
FT      Misc-difference 232
FT      /label= Thr, Ser
FT      Misc-difference 236
FT      /label= Asp, Ser
XX      DE19804210-A1.
XX      12-AUG-1999.
XX
XX      03-FEB-1998; 98DE-1004210.
XX
XX      03-FEB-1998; 98DE-1004210.
XX
XX      (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
XX      Morris P, Stiefel T, Voelter W, Walters P;
XX      WPI; 1999-445335/38.
XX
XX      Preparation of mistletoe lectins in heterologous systems,
XX      particularly for use as anticancer agents and immunostimulants
XX
XX      Claim 2; Page 27; 78pp; German.
XX
XX      This invention describes a novel mistletoe lectin (I) and its fragments
XX      which have antitumour and immunostimulatory activity. The A-chain (MLA)
XX      ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX      lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX      fragments are used to treat uncontrolled cell growth (particularly
XX      cancers) and if they lack cytotoxicity, to increase the strength of the
XX      immune response, particularly to a co-administered antigen
XX      (tumour-associated, bacterial or viral). The method allows production of
XX      mistletoe lectin, and its individual chains, in many different isoforms
XX      and on a large scale, at any time of the year. Recombinant products are
XX      free from toxins present in natural mistletoe extracts. This sequence
XX      represents a consensus sequence of the mistletoe lectin A chain (MLA)
XX      described in the invention.
XX
XX      Sequence 255 AA;

```

```

alignment_scores:
  Quality: 599.00      Length: 140
  Ratio: 4.538        Gaps: 0
  Percent Similarity: 94.286      Percent Identity: 84.286

```

alignment\_block:

US-09-627-165B-15 x AAY25971 ..

Align seg 1/1 to: AAY25971 from: 1 to: 255

```

1  TACACAGATCTGAGGATACGGCGGATAGGACCAAGATCCCTGGG 50
116 Tyr***AspLeuGlnArgTyrAlaGlyHisArgAspGlnIleProLeuGln 132
51  TATAGAGAACTCATTCATCCGTCGCGGCTTCGTTATCCAGCGGGGA 100
132 yile**GlnLeuIleGlnSerVal**AlaLeuArg***ProGlyGlyS 149
101 GCACCGGGGCCAAGCTGTTCCCTTATATCCCTCATTCAGATGATCC 150
149 erThrArg**GlnAlaArgSerIleLeuIleGlnMetIleSer 165
151 GAGCGCGGAGATTCATCCATCTTTGGAGGCGCGCATTAATTA 200
166 GluAlaAlaArgPheAsnProIleLeuTyrPArg**ArgGln**IleAs 182

```

```

201 CAGCGGAGGATCATTTCTCCCGACATGATACATGCTGACCTGGAGACTA 250
182 nserGly**SerPheLeuProAsp***TyrMetLeuGlnLeuGlnThrS 199
251 GTTGGGGCCACAAATCCACGCAATGCTACGAGTGTGGCGCTTTT 300
199 erTTPGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 215
301 AATAACCATTTGCGTTGGGTATATCCACCGGTAACTTCGTGACGTTGAG 350
216 AsnAsnPro**ArgLeuAlaIle****GlyAsnPheValThrLeu** 232
351 CAATGTTCCGCGACGTGATCCGACGCTTAGCGATCATGTTGTATGTA 400
232 *AsnValArg**ValIleAlaSerLeuAlaIleMetLeuPheValCysG 249
401 GGCACCGACCATCTCTCCTCC 420
249 LysIleArgProSerSerSer 255

```

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqp\_emb1/AA1999.DAT:AA25974

seq\_documentation\_block:

ID AAY25974 standard; protein: 255 AA.

AC AAY25974;

DT 18-OCT-1999 (first entry)

DE Mistletoe lectin A chain MLA consensus protein sequence 2.

KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;  
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;  
 KW cancer; cytotoxicity; antigen; isoform.

OS Viscum album.

Key Location/Qualifiers

```

FT      Misc-difference 15
FT      /label= Asp, Glu
FT      Misc-difference 63
FT      /label= Gly, Gln
FT      Misc-difference 66
FT      /label= Ile, Val
FT      Misc-difference 75
FT      /label= Leu, Ala
FT      Misc-difference 107
FT      /label= Asp, Arg, none
FT      Misc-difference 113
FT      /label= Asn, Thr
FT      Misc-difference 117
FT      /label= Pro, Thr
FT      Misc-difference 134
FT      /label= Asp, Glu
FT      Misc-difference 141
FT      /label= Ser, Thr
FT      Misc-difference 145
FT      /label= Phe, Tyr
FT      Misc-difference 152
FT      /label= Thr, Ala
FT      Misc-difference 177
FT      /label= Ala, Tyr
FT      Misc-difference 180
FT      /label= Tyr, Asp
FT      Misc-difference 185
FT      /label= Ala, Glu
FT      Misc-difference 191
FT      /label= Val, Met
FT      Misc-difference 219
FT      /label= Ile, Phe
FT      Misc-difference 224
FT      /label= Phe, Ser

```

```
FT Misc-difference 225
FT /label= Pro, Thr
FT Misc-difference 232
FT /label= Thr, Ser
FT Misc-difference 236
FT /label= Asp, Ser
FT
FT DE19804210-A1.
PN
XX 12-AUG-1999.
XX
XX 03-FEB-1998: 98DE-1004210.
XX
XX 03-FEB-1998: 98DE-1004210.
XX
XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
XX Morris P, Stiefel T, Voelter W, Walters P.
XX
XX WPI: 1999-445335/38.
XX
XX Preparation of mistletoe lectins in heterologous systems,
XX particularly for use as anticancer agents and immunostimulants
XX
XX Claim 5: Page 30: 78pp: German.
XX
XX This invention describes a novel mistletoe lectin (I) and its fragments
XX which have antitumour and immunostimulatory activity. The A-chain (MLA)
XX of the mistletoe lectin binds to, and inactivates, the 28S subunit of
XX ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX fragments are used to treat uncontrolled cell growth (particularly of the
XX cancers) and if they lack cytotoxicity, to increase the strength of the
XX immune response, particularly to a co-administered antigen
XX (tumour-associated, bacterial or viral). The method allows production of
XX mistletoe lectin, and its individual chains, in many different isoforms
XX and on a large scale, at any time of the year. Recombinant products are
XX free from toxins present in natural mistletoe extracts. This sequence
XX represents a consensus sequence of the mistletoe lectin A chain (MLA)
XX described in the invention.
XX
XX Sequence 255 AA:
XX
XX alignment_scores:          length: 140
XX      quality: 599.00      gaps: 0
XX      ratio: 4.538
XX Percent Similarity: 94.286 Percent Identity: 84.286
XX
XX alignment_block:
XX US-09-627-165B-15 x MAY25974 ..
XX
XX Align seq 1/1 to: MAY25974 from: 1 to: 255
XX
XX 1 TACACAGATCTGAGCGATACGCGGTATAGGAGCCAGATCCCTCTGG 50
XX |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 116 Tyr***AspLeuGluArgTyrAlaGlyHisArgAspGlnIleProLeuG 132
XX 51 TATFAGAGCACTCATTCATCCGCTGCGGCTCGTTATCCGATCCAGCGGCA 100
XX |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
XX 132 ylle***GlnLeuIleGlnSerVal***AlaLeuArg***ProGlyGlys 149
XX 101 GCACCGGCGCCAGCTGTTCCCTTATATATCTCATTCAGATGATCTCC 150
XX |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
XX 149 erThrArg***GlnAlaArgSerIleLeuIleLeuIleGlnMetIleSer 165
XX 151 GAGCGCGGAGATCATTCATCCATCTTTGGAGGGGCTGCGCAATCATTA 200
XX |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
XX 166 GlnAlaAlaArgPheAsnProIleLeuTyrArg***ArgGln***IleAs 182
XX 201 CAGCGGGAGATCATTTCTTCCGACATGTACATGCTCGAGCTGAGACTA 250
XX |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
XX 182 nserGly***SerPheLeuProAsp***TyrMetLeuGlnLeuGlnThr 199
```

```
251 GTTGGGCCAACAATCCAGCAGATCCAGAGTCTACGGATGCGCTTTT 300
XX |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
XX 199 erTyrGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 215
XX 301 AATAACCATTTGGTGGGTATATCCACCGGTATACCTTGCTGACGTGAG 350
XX |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
XX 216 AsnAsnPro***ArgLeuAlaIle*****GlyAsnPheValThrLeu** 232
XX 351 CAATGTTCCGACGCTGATCCGACGCTTACGATCATGTTGTATGATGA 400
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
XX 232 *AsnValArg***ValIleAlaSerLeuAlaIleMetLeuPheValCysG 249
XX 401 GGCACCGACATCTTCCTCC 420
XX :|||||:|||||:|||||
XX 249 yGluArgProSerSerSer 255
XX
XX seq_name: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA199.DAT:AAV25977
XX
XX seq_documentation_block:
XX ID AAV25977 standard; protein: 255 AA.
XX
XX AC AAV25977;
XX
XX DI 18-OCT-1999 (first entry)
XX
XX DE Mistletoe lectin A chain MLA consensus protein sequence 3.
XX
XX KW Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
XX ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
XX lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX cancer; cytotoxicity; antigen; isoform.
XX
XX OS Viscum album.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 15
XX FT /label= Asp, Glu
XX FT Misc-difference 63
XX FT /label= Gly, Gln
XX FT Misc-difference 66
XX FT /label= Ile, Val
XX FT Misc-difference 75
XX FT /label= Leu, Ala
XX FT Misc-difference 107
XX FT /label= Asp, none
XX FT Misc-difference 113
XX FT /label= Asn, Thr
XX FT Misc-difference 117
XX FT /label= Pro, Thr
XX FT Misc-difference 134
XX FT /label= Asp, Glu
XX FT Misc-difference 141
XX FT /label= Ser, Thr
XX FT Misc-difference 145
XX FT /label= Phe, Tyr
XX FT Misc-difference 152
XX FT /label= Thr, Ala
XX FT Misc-difference 177
XX FT /label= Ala, Tyr
XX FT Misc-difference 180
XX FT /label= Tyr, Asp
XX FT Misc-difference 185
XX FT /label= Ala, Glu
XX FT Misc-difference 191
XX FT /label= Val, Met
XX FT Misc-difference 219
XX FT /label= Ile, Phe
XX FT Misc-difference 224
XX FT /label= Pro, Ser
XX FT Misc-difference 225
XX FT /label= Pro, Thr
XX FT Misc-difference 232
```



alignment\_scores:                      Length:    122  
                     Quality:    408.50                      Gaps:    1  
                     Ratio:    4.045  
 Percent Similarity:    82.787                      Percent Identity:    69.672

alignment\_block:

US-09-627-165b-15 x AAB47096    ..

Align seg 1/1 to: AAB47096 from: 1 to: 120

```

157 GCGAGATTCCATCCCATCTTTGGAGGGCTGCCCATATCACTTAACAGCGG 206
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1 AlaArgPheAnpProIle**TrpArgLeuAlrArgGlnIleAsnSerG1 17
207 GGAGCATTTCTTCCGACATGATCGTGCAGCTGCAGAGATGATGGG 256
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
17 yGlnSerSerProProAsnMetIlyrMetLeuGlnLeuGlnThrSerTrpG 34
257 GCCAACAATCCAGCAGTCCAGCAGTCTACGATGCGGCTTTTAATAC 306
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
34 LyrGlnSerThrGlnValGlnGlnSerLysAspGlyIlePheAsnThr 50
307 CCATTTCGTTGGTATATCCAGCGGTACTTGTGACGTGACCAATGT 356
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
51 GlnIleArgLeuGlnIleSerAlaGlnAsnPheValThr**SerAsnVa 67
357 TCGCGACGTATGCGCAGCTTACGATCATGTGTTGTATGATGAGGAC 406
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
67 lArGAspValIleSerSerLeuAlaIleMetLeuPheGlnLucysSerGlyA 84
407 GACCATCTTCCCTCCGACGTCGCTATGGCCGCTGTCATACAGCCGTC 456
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
84 rGrProPheSerSerLeuAspHisProSerProLeuLeuLeuArgSerVal 100
457 TTGGAATAATAGCGGCGCGTCGACGATGTTACTGACACTGCTTCCGAAC 506
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
101 ValAsp.....AlaAlaAsnAspValThrCysThr**SerGlnThr 114
507 CACCGTCGCGATCGTA 522
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
114 oThValAlrGlnIleVal 119

```

seq\_name: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1987.DAT: AAP70097

seq\_documentation\_block:

ID AAP70097 standard; protein; 332 AA.

```

XX AAP70097;
AC
XX
XX 09-APR-1991 (first entry)
DE
XX
XX Ricin A.
KM
XX Ricin A; Met-aminopeptidase.
OS
XX Escherichia coli.
PN
XX EP219237-A.
PD
XX 22-APR-1987.
PF
XX 19-SEP-1986; 86EP-0307242.
XX
XX 06-MAY-1986; 86US-0860330.
PR
XX 20-SEP-1985; 85US-0778414.
XX
XX (CERTU ) CETUS CORP.
PA
XX Benbasat A, Bauer KA, Chang S, Chang SY;
PI
XX WPI: 1987-110172/16.
XX
XX DR N-PSDB; AAN70152.
XX

```

PT N-terminal methionine free proteins prodn. - by using host  
 PT transformed with vector to express a methionine-amino-peptidase  
 PS Disclosure; Fig. 4; 20pp; English.

XX Ricin A may be produced in a form which lacks an N-terminal Met  
 CC using Met-aminopeptidase from E.coli.  
 CC

SQ Sequence 332 AA.

alignment\_scores:                      Length:    178  
                     Quality:    259.50                      Gaps:    5  
                     Ratio:    2.257  
 Percent Similarity:    64.607                      Percent Identity:    41.011

alignment\_block:

US-09-627-165b-15 x AAP70097    ..

Align seg 1/1 to: AAP70097 from: 1 to: 332

```

1 TACACAGATCTGGAGACGATACCGGTCAT...AGGACGAGATCCCTCT 47
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
158 TyrAspArgLeuGlnGlnLeuAlaGlnAsnLeuArgGlnAsnIleGluLe 174
48 GGGTATAGAGAACTCATTCATCCCTCTCGCGCTTCGTTAT..... 90
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
174 uGlnAsnGlyProLeuGlnGlnAlaIleSerAlaLeuLyrTrpTrpSerT 191
91 ..CCAGCGCGCAGCAACCGCGCCCAAGCTCGTCCCTTAATATCCTCAT 138
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
191 hrcIyGlyThrGlnLeuProThrLeuAlaIleArgSerPheIleIleCysIle 207
139 CAGATATCTCCGAGCGCCGAGATTCATCCATCTTTGGAGGCTCG 188
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
208 GlnMetIleSerGlnAlaAlaIleArgPheGlnTrpIleGlnGlnMetAr 224
189 CCATTCATTTAACAGCGGGGAGATCTTCTCCGACATGTCATGCTCG 238
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
224 gThrArgIleArgTrpAsnArgTrpSerAlaProAspProSerValIleT 241
239 AGCTGACGATCTGGAGCGCCCAACATCCAGCAGCAAGTCCAGAGCTTAC 288
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
241 hrlLeuGlnAsnSerTrpGlyArgLeuSerThrAlaIleGlnIleSerAsn 257
289 GATGCGCTTTTAAATACCATTTGGTGGTATATCCAGCGTAACTT 338
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
258 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnIleSerly 274
339 CGTACGCTTGACAAATGTCGACGATGATGCCAGCTTAGCGATCATCT 388
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
274 sPheSerValTrpAspValSerIleLeuIleProIleIleAlaLeuMetV 291
389 TGTGTATGTAGGACCGACCATCTTCCCTCCGACGTCGCGCTTGGCGG 438
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
291 alTYrArgCysAlaProProProSerSerGln.....PheSer 303
439 CTGCTCATACGACCGCGCTTGAATAATAGCGGCGCTGCAGCATGTTAC 488
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
304 LeuLeuIleArgProValValPheAsnPheAsnAla.....AspVal... 317
489 CTGCACGCTTCCGAACCCAGCGTGCGCATCGTA 522
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
318 .CysMetAspProGlnProIleValAlrGlnIleVal 328

```

seq\_name: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1987.DAT: AAP70838

seq\_documentation\_block:

ID AAP70838 standard; protein; 332 AA.

```

XX AAP70838;
AC
XX
XX 18-FEB-1991 (first entry)
XX

```

Quality:	259.50	Length:	178
Ratio:	2.257	Gaps:	5
Percent Similarity:	64.607	Percent Identity:	41.011

US-09-627-165B-15 x AAP70838

Align seg 1/1 to: AAP70838 from: 1 to: 332

1 TACACATCTTTGGAGGCGATCGGGCTCAT...AGGACACGATCCCTCT 47  
 158 Tyrsparfyeuglucglnleuaslactylasylasleuafglsuaslleglule 174  
 48 GCGTATAGAGGACGATTCATCCGCTCTCGGCGCTCTGTTAT..... 90  
 174 ugllyasnglpyroleuclucinalleesrlaleuurytyrtyrserr 191

```

91  ..CCAGGGGGGAGCAGCCCGGGCCCAAGCTCGTCCCTATTAATCTCATTT 138
191 hnglygylthnglnneuprothleuallargserPheIleIleCysIle 207
139 CAGATGATCTCCGAGGCGCGGAGATTCAATCCCATCTTTTGTGAGGGCTCG 188
208 glnmetIleSerGlnuaIaalaIargPhegInTyrlleGlnuglyumetIar 224
189  CCATACATTAAACGAGGGGAGTCATTTCTTCCGACATGTACATGCTCG 238
224  gThArgIleArgTyrrasnIarGrserralaprasProserValIleIar 241
239 AGCTGGAGACTAGTTGGGGCCCAACATCCGACGAAGTCAGCATGCTACG 288
241 hrleuGlnuaInserTPGlyIargLeuSerThrIaIleGlnGlnIuserasn 257
289 GATGGCGCTTTTAAATACCATTTGGTGGGTATATCCACCGGTAACTT 338
258 GlnclYalalPhealaserProIleGlnleuInIarGargasnGlyserIy 274
339 CGTGGACGTTGAGCAATGTTCCGAGAGTGATGGCCAGATGAGCATGATCT 388
274 sPheSerValTyrrasPalserIleleuIleproIlellealIleuMetY 291
389 TGGTTTATGTAAGGGAGCGACACATCTTCCTCCGACGTGGCTATGGCGG 438
291 alTyrrGlyCysalalProProProserSerGln.....PheSer 303
439 CTGTGTCATACGACCGGCTTTGGAAAAATAGCGGCGCGTACAGATGTAC 488
304 leuIleuIleArgProValIalProasnPheasnala.....AspVal.. 317
489 CTGCACTGCTCCGACACCGACCGCTGGCGCATGCA 522
318 ..CysMetIasProGlnProIleValIargIleVal 328

```



Wed Jul 31 08:46:42 2002

OM of: US-09-627-165B-15 to: Issued\_Patents\_AA: \* out\_format : pfs  
Date: Jul 30, 2002 3:54 PM

About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+npz.model -DEV=xlp  
-O=/cg2\_1/USPTO.spool/US09627165/Runat\_30072002\_151138\_3219/app\_query.fasta\_1.1012  
-DB=Issued\_Patents\_AA -QFMT=fastan -SUFFIX=ra1 -GAPOP=12.000  
-GAPBX=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blodum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORML=HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09627165.@CGN1.1.44 -NCPU=6 -ICPU=3 -LONLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLIFY -WAIT -THRBADS=1

Search information block:

Query: US-09-627-165B-15  
Query length: 522  
Database: Issued\_Patents\_AA: \*  
Database sequences: 231628  
Database length: 24425594  
Search time (sec): 39.330000

Score list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation	
/cg2_6/p/ptodata/2/1aa/5B.COMB.pep:US-08-776-059-39 +			619.00	1330.02	4.4e-68	235
/cg2_6/p/ptodata/2/1aa/5B.COMB.pep:US-08-776-059-31 +			619.00	1349.14	4.0e-68	253
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-378-761A-77 +			256.00	533.96	5.5e-23	540
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-485-286-77 +			256.00	533.96	5.5e-23	540
/cg2_6/p/ptodata/2/1aa/5B.COMB.pep:US-08-378-761A-71 +			242.50	504.12	2.6e-21	534
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-485-286-71 +			235.50	497.58	1.3e-20	250
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-07-988-430-1 +			217.00	455.72	2.5e-18	267
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-218-303-16 +			217.00	455.72	2.5e-18	267
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-425-336-1 +			217.00	455.72	2.5e-18	267
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-488-113B-1 +			217.00	455.72	2.5e-18	267
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-477-484B-1 +			217.00	455.72	2.5e-18	267
/cg2_6/p/ptodata/2/1aa/5B.COMB.pep:US-08-646-360-1 +			217.00	455.72	2.5e-18	267
/cg2_6/p/ptodata/2/1aa/5B.COMB.pep:US-08-338-793D-61 +			217.00	455.72	2.5e-18	267
/cg2_6/p/ptodata/2/1aa/5B.COMB.pep:US-08-839-765-1 +			217.00	455.72	2.5e-18	267
/cg2_6/p/ptodata/2/1aa/5B.COMB.pep:US-09-136-389-1 +			217.00	455.72	2.5e-18	267
/cg2_6/p/ptodata/2/1aa/5B.COMB.pep:US-08-356-786-8 +			213.00	445.85	8.3e-18	268
/cg2_6/p/ptodata/2/1aa/5B.COMB.pep:US-08-378-761A-77 +			213.00	445.85	8.3e-18	268
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-324-301-15 +			166.50	342.91	4.6e-12	282
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-378-761A-74 +			163.50	336.90	1.1e-11	267
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-485-286-74 +			156.50	321.62	7.6e-11	261
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-485-286-79 +			156.50	320.41	8.0e-11	289
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-07-988-430-7 +			154.50	317.09	1.4e-10	263
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-07-988-430-7 +			154.50	317.09	1.4e-10	263
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-425-336-7 +			154.50	317.09	1.4e-10	263
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-477-484B-7 +			154.50	317.09	1.4e-10	263
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-488-113B-7 +			154.50	317.09	1.4e-10	263
/cg2_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-646-360-7 +			154.50	317.09	1.4e-10	263
/cg2_6/p/ptodata/2/1aa/5B.COMB.pep:US-08-839-765-7 +			154.50	317.09	1.4e-10	263
/cg2_6/p/ptodata/2/1aa/5B.COMB.pep:US-09-136-389-7 +			154.50	317.09	1.4e-10	263

/cg2\_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-488-113B-6 + 153.50 315.62 1.7e-10 247  
/cg2\_6/p/ptodata/2/1aa/5A.COMB.pep:US-08-477-484B-6 + 153.50 315.62 1.7e-10 247  
/cg2\_6/p/ptodata/2/1aa/5B.COMB.pep:US-08-646-360-6 + 153.50 315.62 1.7e-10 247  
seq\_name: /cg2\_6/p/ptodata/2/1aa/5B.COMB.pep:US-08-776-059-39

seq\_documentation\_block:

Sequence 39, Application US/08776059B  
Patent No. 6271368  
GENERAL INFORMATION:  
APPLICANT: LENTZEN, Hans  
APPLICANT: BECK, Jurgen  
APPLICANT: BAUR, Axel  
APPLICANT: ZINKE, Holger  
TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)  
FILE REFERENCE: 674503-2003  
CURRENT APPLICATION NUMBER: US/08/776, 059B  
CURRENT FILING DATE: 1999-06-19  
EARLIER APPLICATION NUMBER: PCT/EP96/02273  
EARLIER FILING DATE: 1996-06-25  
EARLIER APPLICATION NUMBER: 95109949.8  
EARLIER FILING DATE: 1995-06-26  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: Patentlin Ver. 2.0  
SEQ ID NO 39  
TYPE: PRT  
ORGANISM: Viscum album  
US-08-776-059-39

alignment\_scores:

Quality: 619.00 Length: 138  
Ratio: 4.654 Gaps: 0  
Percent Similarity: 96.377 Percent Identity: 86.232

alignment\_block:

US-09-627-165B-15 x US-08-776-059-39  
Align seg 1/1 to: US-08-776-059-39 from: 1 to: 235

1 TACACAGATCTGAGAGATACCGCGCTATAGGACGACATCCCTCGG 50  
|||||  
98 TTPPAGAGACTGATTCATCCGCTCGGCGCTGATTCAGCGGCA 100  
|||||  
114 YLEASPGINLEULIEGLINSEVALTHRLAULNARPHROGLYGLYS 131  
|||||  
101 GCACCGCGCGCCAGCTGCTTCCTATATATCTCATTCAGATGATCC 150  
|||||  
131 ETRHARATHRGINALNARGLSERILEULEULEULEULEULEULE 147  
|||||  
151 GAGCGCGCGAGATTCATCCATCTTTGGAGGCGCGCATATACATTA 200  
|||||  
148 GLULALALNARPHENAPROILLEUTTPAGALNARGLNITLYLLEAS 164  
|||||  
201 CAGCGGAGATCATTTCTCCGACATGATCATGCTGAGCTGAGACTA 250  
|||||  
164 NSEGLYALSERHEULNARPHROASPAVALRYMETLEUGLULEGLUTHS 181  
|||||  
251 GTTGGGCGCAACATTCACCGCAAGTCCAGAGCTCAGGATGGCGTTT 300  
|||||  
181 ETRFGLGLNGLINSETHRGLNVALINLISSETHRAPHPELYALPHE 197  
|||||  
301 AATPACCCATTTGGTGGTATATCCACCGGTAACCTGCTGACGTTGAC 350  
|||||  
198 ASHAPNPROILLEAULNARPHROGLYASNHVALTHRLNLEUTH 214  
|||||  
351 CAATGTCGCGAGATTCAGCGGATTCAGATTCAGATTCAGATTCAG 400  
|||||  
214 RANVALNARPHASPAVALLEALNARPHROGLYASNHVALTHRLNLEUTH 214

Wed Jul 31 08:46:42 2002

us-09-627-165b-15.ra1

Page 2

401 GGGAGCGACCATCT 414  
231 YGLIARPRoser 235

seq\_name: /cgn2\_6/plodata/2/1aa/5A.COMB.pep:US-08-776-059-31

seq\_documentation\_block:

; Sequence 31, Application US/08776059B  
; Patent No. 6271368  
; GENERAL INFORMATION:  
; APPLICANT: LENTZEN, Hans  
; APPLICANT: ECK, Jurgen  
; APPLICANT: BAUR, Axel  
; APPLICANT: ZINK, Holger  
; TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)  
; FILE REFERENCE: 674503-2003  
; CURRENT APPLICATION NUMBER: US/08/776,059B  
; EARLIER FILING DATE: 1999-06-19  
; EARLIER APPLICATION NUMBER: PCT/EP96/02273  
; EARLIER FILING DATE: 1996-06-25  
; EARLIER APPLICATION NUMBER: 95109949.8  
; NUMBER OF FILING DATE: 1995-06-26  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Patentn Ver. 2.0  
; SEQ ID NO 31  
; TYPE: PRT  
; LENGTH: 253  
; ORGANISM: Viscum album  
US-08-776-059-31

alignment\_scores:

Quality: 619.00 Length: 138  
Ratio: 4.654 Gaps: 0  
Percent Similarity: 96.377 Percent Identity: 86.232

alignment\_block:

US-09-627-165b-15 x US-08-776-059-31 ..

Align seg 1/1 to: US-08-776-059-31 from: 1 to: 253

1 TACACAGATCTGGAGCATACCGGATAGGAGGACGATCCCTGGG 50  
116 TyrProbspleugluariglyralaglyHlaarghsplnleproleu 132  
51 TATAGAGAACTCATTCATCCGCTCTGGGCTTCGTATCCAGGGGGA 100  
132 YLleaspGlnleuileglnserValtrAlaleuargpneproglylys 149  
101 GCACCCGGGCGCAAGCTGCTCCCTTATATCTCATTCAGATCTCC 150  
149 erThraTgtrGlnAlaargSerleleuileuileglnMetIleSer 165  
151 GAGCGCGGATTCATCCATCTTTGGAGGCTGCCAATACATTA 200  
166 GlnAlaAlaarghsnproleleuTprghlaarglInIryIleS 182  
201 CACCGGGAGTCTTCTCCGAGATGATCTCTGAGCTGAGACGA 250  
182 nserGlyAlaSerPheleuProaspValTyrMetleuGlnleuGly 199  
251 GTTGGGGCCAAATCCACCAAGTCCAGAGTCTACGATGGCGTTT 300  
199 erTprgIynglnserThrGlnValGlnHisSerThrAspGlyValPhe 215  
301 AATTAACCATTTGCTGGGTATATCCAGGCTTAACTCTGAGCTTGA 350  
216 AsnAsnProleleargleuAlaIleproproGlysnPheValThrleu 232  
351 CATGTTCGGGAGCTGATGCCAGGTTTAGCATCATCTGTTGTATGTA 400  
232 TAsnValArgAspValIleAlaSerleuAlaIleMetleuPheValCy 249

401 GGGAGCGACCATCT 414  
249 YGLIARPRoser 253

seq\_name: /cgn2\_6/plodata/2/1aa/5A.COMB.pep:US-08-378-761A-77

seq\_documentation\_block:

; Sequence 77, Application US/08378761A  
; Patent No. 5635384  
; GENERAL INFORMATION:  
; APPLICANT: MALSH, TERENCE A  
; APPLICANT: HEY, TIMOTHY D  
; APPLICANT: MORGAN, ALICE ER  
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ANDREA T. BORUCKI  
; STREET: 9330 ZIONSVILLE ROAD  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: US  
; ZIP: 46268  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/378,761A  
; FILING DATE: 26-JAN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BORUCKI, ANDREA T  
; REGISTRATION NUMBER: 33651  
; REFERENCE/DOCKET NUMBER: 38272B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317) 337-4846  
; INFORMATION FOR SEQ ID NO: 77:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 540 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-378-761A-77

alignment\_scores:

Quality: 256.00 Length: 177  
Ratio: 2.246 Gaps: 4  
Percent Similarity: 64.407 Percent Identity: 39.548

alignment\_block:

US-09-627-165b-15 x US-08-378-761A-77 ..

Align seg 1/1 to: US-08-378-761A-77 from: 1 to: 540

1 TACACAGATCTGGAGCATACCGGATAGGAGGACGATCCCTGGG 50  
123 TyrAspArgleuglnleuGlnleuArgGlnleuGlnleuGln 139  
51 TATAGAGAACTCATTCATCCGCTCTGGGCTTCGTATCCATTC 91  
139 YThrGlyProleuGlnAspAlaIleSerAlaIleuTyrTyrSerThr 156  
92 CAGGCGGAGCAACCGCGGACGATCGTTCCTTATATCCATTCAG 141  
156 YsgIythrGlnIleProThrleuAlaargSerPheMetValCysIle 172  
142 ATGATCTCCGAGGCGGAGATTCATCCATCTTTGGAGGCGCTGCC 191

```

|||||
173 MetIleSerGluAlaAlaArgPheGlnTyrIleGluGlyGluMetArgTh 189
192 ATACATTAAACAGCGGGAGCTATTCTTCCCGACATGTCATCGCCGAGC 241
189 rArgIleArgTyrAsnArgArgSerAlaProAspProSerValIleThrL 206
242 TGGAGACTACTGTGGGGCCACATCCAGCAGATCCAGAGTCTACGAT 291
206 eugIuAsnSerTrrpGlyArgLeuSerThrAlaIleGlnGluSerAsnGln 222
292 GCGCTTTTATTAACCATTTGCGTGGGTATATCCACCGCTACTTCGT 341
223 GlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerIysph 239
342 GACGTATGACAAATGTTCGCCAGCTGATCGCAGCTTACGATGTTGT 391
239 eAsnValTyrAspValSerIleLeuIleProIleIleAlaLeuMetValT 256
392 TTGTATGTAGGACCGACCATCTCTCCGACGTGCGCTATTGGCCGCTG 441
256 YrArgCysAlaProProSerSerGln.....PheSerIleu 268
442 GTCATACGACCCGCTTGGAAATAGCGGGCCGCTCGACGATGTACTCG 491
269 LeuIleArgProValValProAsnPheAsnAla.....AspVal...Cy 282
492 CACTGCTTCCGAACCCAGCGTGCATCGTA 522
282 sMetAspProGluProIleValArgIleVal 292

seq_name: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:us-08-485-286-77

seq_documentation_block:
; Sequence 77, Application US/08485286
; Patent No. 5646026
; General Information:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC Compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378761
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 amino acids

```

```

; TYPE: amino acid
; STRANDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-286-77

alignment_scores:
    Quality: 256.00      Length: 177
    Ratio: 2.246         Gaps: 4
    Percent Similarity: 64.407      Percent Identity: 39.548

alignment_block:
US-09-627-165B-15 x US-08-485-286-77 ..
Align seg 1/1 to: US-08-485-286-77 from: 1 to: 540

1 TACACAGATCTGAGACCGATACGCCGCTCATATGAGACCAATCCCTGGG 50
|||
123 TyrAspArgLeuGlnGluGlnLeuGlyGlyLeuArgGluGlnIleGluLeuG 139
51 TATAGAGAGACTATTCATCCGCTCGCGCTTCTTAT.....C 91
139 YthrGlyProLeuGlnAspAlaIleSerAlaLeuTyrTyrSerThrC 156
92 CAGGCGGACGACACCGGCGCCAGCTGCTTCCCTATATACCTCATTCAG 141
156 YsGlyThrGlnIleProThrLeuAlaArgSerPheMetValCysIleGln 172
142 ATGATCTCCGAGCGCGGATTCATCCATCTTTGGAGGCTCGCA 191
173 MetIleSerGluAlaAlaArgPheGlnTyrIleGluGlyGluMetArgTh 189
192 ATACATTAAACAGCGGGAGCTATTCTTCCGACATGTACATCCTCGAGC 241
189 rArgIleArgTyrAsnArgArgSerAlaProAspProSerValIleThrL 206
242 TGGAGACTACTGTGGGGCCACATCCAGCAGATCCAGCTTACGAT 291
206 eugIuAsnSerTrrpGlyArgLeuSerThrAlaIleGlnGluSerAsnGln 222
292 GCGCTTTTATTAACCATTTGCGTGGGTATATCCACCGTACTTCGT 341
|||
223 GlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerIysph 239
342 GACGTATGACAAATGTTCGCCAGCTGATCGCAGCTTACGATGTTGT 391
239 eAsnValTyrAspValSerIleLeuIleProIleIleAlaLeuMetValT 256
392 TTGTATGTAGGACCGACCATCTCTCCGACGTGCGCTATTGGCCGCTG 441
256 YrArgCysAlaProProSerSerGln.....PheSerIleu 268
442 GTCATACGACCCGCTTGGAAATAGCGGGCCGCTCGACGATGTACTCG 491
269 LeuIleArgProValValProAsnPheAsnAla.....AspVal...Cy 282
492 CACTGCTTCCGAACCCAGCGTGCATCGTA 522
282 sMetAspProGluProIleValArgIleVal 292

seq_name: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:us-08-356-786-10

seq_documentation_block:
; Sequence 10, Application US/08356786
; Patent No. 5877305
; General Information:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker

```



```

113 TyrGlyAspLeuGlnArgTyrPheAlaHisGlnSerArgGlnGlnIleProLe 129
48 GGGATATAGAGAACTCATTCATCCGCTCGCGCTTCGTATCCAGGG 97
129 uGlyLeuGlnAlaLeuThrHisGlyIleSerPhePheArgSerGlyGlyA 146
98 GCAGCACCAGCCGCGCCAGAGTCGCTCCCTTATATCCATCCATTCAGATGAG 147
146 snAspAsnGlnGluLysAlaArgPheArgTyrIleSerAsnArgValAlaSerI 162
148 TCCGAGCGCCGCGAGATTCATCCCATCTTTTGACGGCTCGCCCAATACAT 197
163 AlaGlnAlaAlaArgPheArgTyrIleSerAsnArgValAlaSerI 179
198 TAAACAGCGGGAGTCATTTCTCCGACATGATGATGCTGAGCTGAGA 247
179 eGlnThrGlyThrAlaPheGlnProAspAlaAlaMetIleSerLeuGlnA 196
248 CTAGTTGGGGCCCAACATCCACGCCAGCTCAGATGAGGCGT 297
196 snAsnTrp...AspAsnLeuArgGlyValGlnGlnSerValGlnAspThr 211
298 TTTAATAACCATTCGCTGGTGGTATATCCACCGCTAACTTCGTGACCT 347
212 PheProAsnGln.....ValThrLe 218
348 GAGCAATGTCGCGAC.....G 364
218 uThrAsnIleArgAsnGlnProValIleValAspSerLeuSerHisProt 235
365 TGATCGCCAGCTTACGATCATCTGTTGTATGT 399
235 hrValAlaValIleuAlaLeuMetLeuPheValCys 246

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-485-286-71

seq_documentation_block:
; Sequence 71, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378761
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-286-71

alignment_scores:
Quality: 235.50 Length: 145
Ratio: 2.403 Gaps: 4
Percent Similarity: 67.586 Percent Identity: 40.000

alignment_block:
US-09-627-165B-15 x US-08-485-286-71 ..
Align seg 1/1 to: US-08-485-286-71 from: 1 to: 250

1 TACACAGATCTGAGCGATGACGCGGTCAAT...AGGACACGATCCCTCT 47
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 TyrGlyAspLeuGlnArgTyrPheAlaHisGlnSerArgGlnGlnIleProLe 129
48 GGGATATAGAGAACTCATTCATCCGCTCGCGCTTCGTATCCAGGG 97
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
129 uGlyLeuGlnAlaLeuThrHisGlyIleSerPhePheArgSerGlyGlyA 146
98 GCAGCACCAGCCGCGCCAGAGTCGCTCCCTTATATCCATCCATTCAGATGAG 147
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
146 snAspAsnGlnGluLysAlaArgPheArgTyrIleSerAsnArgValAlaSerI 162
148 TCCGAGCGCCGCGAGATTCATCCCATCTTTTGACGGCTCGCCCAATACAT 197
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 AlaGlnAlaAlaArgPheArgTyrIleSerAsnArgValAlaSerI 179
198 TAAACAGCGGGAGTCATTTCTCCGACATGATGATGCTGAGCTGAGA 247
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 eGlnThrGlyThrAlaPheGlnProAspAlaAlaMetIleSerLeuGlnA 196
248 CTAGTTGGGGCCCAACATCCACGCCAGCTCAGATGAGGCGT 297
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 snAsnTrp...AspAsnLeuArgGlyValGlnGlnSerValGlnAspThr 211
298 TTTAATAACCATTCGCTGGTGGTATATCCACCGCTAACTTCGTGACCT 347
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
212 PheProAsnGln.....ValThrLe 218
348 GAGCAATGTCGCGAC.....G 364
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
218 uThrAsnIleArgAsnGlnProValIleValAspSerLeuSerHisProt 235
365 TGATCGCCAGCTTACGATCATCTGTTGTATGT 399
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
235 hrValAlaValIleuAlaLeuMetLeuPheValCys 246

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-07-901-707-1

seq_documentation_block:
; Sequence 1, Application US/07901707
; Patent No. 5376546
; GENERAL INFORMATION:
; APPLICANT: Bernhard, Susan L.
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Steve F.
; APPLICANT: Lane, Julie A.
; TITLE OF INVENTION: Materials Comprising and Methods of
; TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell

```





Page 8

```

339 CGTAGCCTTGAGCATTTGGCAGCATGTCAGCAGTCATCAT 388
      :          :          :          :          :
239 sphaSeValTYtAspValSerIleuIleIleProIleIleAlaValuMeCy 256
      :          :          :          :          :
389 TGTTTGATGTAGGACCGACCATCTTCC 417
      :          :          :          :          :
256 aLYtArGcYsAlaProProPseSer 265
      :          :          :          :          :

seq_name: /cgn2_6/plodata/2/1aa/5A_COMB pep:US-08-488-113B-1

seq_documentation_block:
? Sequence 1, Application US/08488113B
? Patent No. 5744580
?
? GENERAL INFORMATION:
? APPLICANT: Better, Marc D.
? APPLICANT: Carroll, Stephen F.
? APPLICANT: Studlika, Gary M.
? TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
? TITLE OF INVENTION: Proteins
? NUMBER OF SEQUENCES: 169
?
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: McAndrews, Held & Malloy, Ltd.
? STREET: 500 West Madison Street, 34th floor
? CITY: Chicago
? STATE: Illinois
? COUNTRY: USA
? ZIP: 60661
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn release #1.0, Version #1.25
? CURRENT APPLICATION DATA: 08/488-113B

```

CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PUBLICATION DATE:

APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/988,430

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER: US 07/901,707
3      FILING DATE: 19-JUN-1992
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER: US 07/787,567
6      FILING DATE: 04-NOV-1991
7      ATTORNEY/AGENT INFORMATION:
8      NAME: MNICHOLAS, Janet M.
9      REGISTRATION NUMBER: 32,918
10     REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C3A#
11     TELECOMMUNICATION INFORMATION:
12     TELEPHONE: 312/707-8889
13     TELEFAX: 312/707-9155
14     TELE: 650 388-1248
15     INFORMATION FOR SEQ ID NO.: 1:
16     SEQUENCE CHARACTERISTICS:
17     LENGTH: 267 amino acids
18     TYPE: amino acid
19     TOPOLOGY: linear
20     MOLECULE TYPE: protein
21     US-08-488-1138-1
22
23 alignment_scores:
24     Quality: 217.00      Length: 143
25     Ratio: 2.399      Gaps: 2
26     Percent Similarity: 65.734      Percent Identity: 39.860

```



## alignment block:

US-09-627-165B-15 x US-08-488-113B-1 ..

Align seg 1/1 to: US-08-488-113B-1 from: 1 to: 267

```
1 TACACAGATCTGGAGCGATACGCCGCTCAT...AGGACCAATCCCTCT 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
123 TyrAspArgLeuGlnLeuAlaGlnLeuAlaArgGlnAsnIleGluLe 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 GCGATATAGAGAACTCATTCATCCGCTCGCGCTTCCTAT..... 90
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 uGlyAsnGlyProLeuGlnLeuAlaIleSerAlaLeuTyrTyrSerT 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
91 ..CCAGCGCGAGCACCGCGGCCCAAGCTCTGCTTCCTTATATCCCAT 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 hrcGlyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 CAGATGATCTCGAGCGCGCGAGATTCATCCCATTCCTTATATCCCAT 188
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
173 GlmMetIleSerGlnAlaAlaArgPheGlnTyrIleGlnGlyGluMetAr 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 CCAATACATTAAACAGCGGGAGTCATTTCTTCCGACATGATGATGCTG 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 gThrArgIleArgTyrAsnArgTyrSerAlaProAspProSerValIleT 206
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 AGCTGAGACTAGTGGGGCCCAACAATCCAGCAAGTCACAGCTACG 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 hTleuGlnAsnSerTpgIlyArgLeuSerThrAlaIleGlnGluSerAsn 222
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
289 GATGCGCTTTTAAATACCATTCCTGCTGGTATATCCACCGCTACT 338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerLy 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
339 CGTGACGTCATGATCTCGGAGCTGATCCGACGCTTAGCATGATCAT 388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 sPheSerValTyrAspAlaSerIleLeuIleProIleIleAlaLeuMetV 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
389 TGTGTATGTAGGAGCGACGACATCTTC 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
256 aTyrArgCysAlaProProSerSer 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

seq_name: /cgn2_6/ptodata/2/1aa/5a_COMB.pep:US-08-477-484B-1

seq_documentation_block:
: Sequence 1, Application US/08477484B
: Patent No. 5756699
: GENERAL INFORMATION:
: APPLICANT: Better, Marc D.
: APPLICANT: Carroll, Stephen F.
: APPLICANT: Studnika, Gary M.
: TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
: TITLE OF INVENTION: Proteins
: NUMBER OF SEQUENCES: 169
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: McAndrews, Held & Malloy, Ltd.
: STREET: 500 West Madison Street, 34th floor
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60661
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DO5
: SOFTWARE: Patentn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA: US/08/477,484B
: APPLICATION NUMBER: US/08/477,484B
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/425,336
: FILING DATE: 18-APR-1995
```

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/938,430

FILING DATE: 09-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-477-484B-1

## alignment\_scores:

Quality: 217.00

Ratio: 2.309

Percent Similarity: 65.734

Percent Identity: 39.860

## alignment\_block:

US-09-627-165B-15 x US-08-477-484B-1 ..

Align seg 1/1 to: US-08-477-484B-1 from: 1 to: 267

```
1 TACACAGATCTGGAGCGATACGCCGCTCAT...AGGACCAATCCCTCT 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
123 TyrAspArgLeuGlnLeuAlaGlnLeuAlaArgGlnAsnIleGluLe 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 GCGTATAGAGAACTCATTCATCCGCTCGCGCTTCCTAT..... 90
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 uGlyAsnGlyProLeuGlnLeuAlaIleSerAlaLeuTyrTyrSerT 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 hrcGlyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 CCAATACATTAAACAGCGGGAGTCATTTCTTCCGACATGATGATGCTG 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 gThrArgIleArgTyrAsnArgTyrSerAlaProAspProSerValIleT 206
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 AGCTGAGACTAGTGGGGCCCAACAATCCAGCAAGTCACAGCTACG 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 hTleuGlnAsnSerTpgIlyArgLeuSerThrAlaIleGlnGluSerAsn 222
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
289 GATGCGCTTTTAAATACCATTCCTGCTGGTATATCCACCGCTACT 338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerLy 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
339 CGTGACGTCATGATCTCGGAGCTGATCCGACGCTTAGCATGATCAT 388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 sPheSerValTyrAspAlaSerIleLeuIleProIleIleAlaLeuMetV 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
389 TGTGTATGTAGGAGCGACGACATCTTC 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```







OM of: US-09-627-165b-15 to: pIR\_71: \* out\_format : pfs  
Date: Jul 30, 2002 3:55 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters: -DEV=api  
-MODEL=frame+ntp.model -DEXT=api  
-O=/cgn2.1/USPTO.spool/US09627165/runtat\_30072002\_151139\_3235/alpha\_query.fasta\_1.1012  
-OB=PIR\_71 -QFMT=fastan -SUFFIX=rrp -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -DELDP=6.000  
-FEAPEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-DEEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pts -NOM=ext -HEPSTLZE=500  
-MINLEN=0 -MAXLEN=2000000000 -USER=US09627165@cgn1.1.172  
-NPU=6 -ICPU=3 -LONGLOG -DEV\_TIMECUT=120 -WARN\_TIMECUT=30  
-NO\_XLPHY -WAIT -THREADS=1

Search information block:  
Query: US-09-627-165b-15  
Query length: 522  
Database: PIR\_71: \*  
Database sequences: 283138  
Database length: 96089334  
Search time (sec): 72.560000

Sequence	Strd	Orig	Zscore	Escore	Len	Document
PIR2:PD0018	+	621.00	1152.96	1.6e-56	254	mistletoe lectin I A chain - Vi
PIR1:TLZLSA	+	283.00	510.91	4.3e-21	528	abrin d precursor - Indian bear
PIR1:TLZLSA	+	259.50	465.94	1.3e-18	576	ricin D precursor - castor bean
PIR2:S53231	+	256.00	459.59	2.9e-18	528	abrin d precursor - Indian bear
PIR2:S16022	+	256.00	459.59	2.9e-18	562	abrin c precursor - Indian bear
PIR1:RLCSAG	+	246.50	449.56	2.9e-17	527	agglutinin precursor - Indian bear
PIR2:S53230	+	246.50	449.56	2.9e-17	527	agglutinin precursor - Indian bear
PIR2:S62627	+	237.00	433.75	2.9e-16	570	agglutinin I precursor - Europe
PIR2:CS39761	+	220.00	393.47	1.5e-14	251	abrin (clone 7.2) precursor - I
PIR2:RLZTPT	+	180.50	333.29	1.7e-10	106	abrin (clone 3.7) precursor - I
PIR1:RLZTPT	+	163.50	279.75	6.0e-08	289	RNA N-glycosidase (EC 3.2.2.22)
PIR2:CS606	+	157.00	279.75	6.0e-08	289	RNA N-glycosidase (EC 3.2.2.22)
PIR2:DE0401	+	156.00	277.71	7.6e-08	261	antiviral protein - Virginal B
PIR2:DE0401	+	156.00	277.71	7.6e-08	294	RNA N-glycosidase (EC 3.2.2.22)
PIR2:CS032	+	154.50	275.52	1.1e-07	247	karsurin-B - Trichosanthes kid
PIR1:RLPUGG	+	154.50	275.52	1.1e-07	286	RNA N-glycosidase (EC 3.2.2.22)
PIR2:U0393	+	152.50	272.76	1.7e-07	247	karsurin - Mongolian snake-ven
PIR2:S17757	+	144.50	255.52	1.2e-06	313	RNA N-glycosidase (EC 3.2.2.22)
PIR2:S55560	+	141.00	249.78	2.8e-06	286	RNA N-glycosidase (EC 3.2.2.22)
PIR2:AS9817	+	140.00	248.17	3.6e-06	278	RNA N-glycosidase (EC 3.2.2.22)
PIR2:JC4840	+	138.00	245.59	5.6e-06	245	RNA N-glycosidase (EC 3.2.2.22)
PIR2:JC4840	+	137.00	241.33	7.4e-06	316	RNA N-glycosidase (EC 3.2.2.22)
PIR2:JC4235	+	136.00	240.39	9.4e-06	286	RNA N-glycosidase (EC 3.2.2.22)
PIR2:JC4235	+	134.50	237.87	1.3e-05	277	RNA N-glycosidase (EC 3.2.2.22)
PIR2:S22494	+	125.00	220.86	0.0001	253	RNA N-glycosidase (EC 3.2.2.22)
PIR2:S22494	+	125.00	217.93	0.0002	283	RNA N-glycosidase (EC 3.2.2.22)
PIR1:RLQHG2	+	124.00	217.64	0.0002	292	beta-luffin - smooth loofah
PIR2:S23519	+	123.50	217.16	0.0002	253	RNA N-glycosidase (EC 3.2.2.22)
PIR2:S28542	+	119.00	209.58	0.0006	253	RNA N-glycosidase (EC 3.2.2.22)
PIR2:S28541	+	119.00	209.58	0.0006	253	RNA N-glycosidase (EC 3.2.2.22)
PIR2:JC4811	+	116.50	204.21	0.0010	272	beta-luffin - beet
PIR2:JUN0108	+	114.00	200.30	0.0018	250	luffin-b - smooth loofah
PIR2:S17519	+	112.00	195.06	0.0031	293	RNA N-glycosidase (EC 3.2.2.22)
PIR2:S17932	+	103.50	181.10	0.0228	236	RNA N-glycosidase (EC 3.2.2.22)
PIR2:AS6923	+	99.00	179.52	0.0683	253	RNA N-glycosidase (EC 3.2.2.22)
PIR2:S01032	+	93.50	159.50	0.2693	318	Shiga-like toxin II chain A pre
PIR2:S21940	+	93.50	159.50	0.2693	319	Shiga-like toxin II chain A - B
PIR2:154695	+	93.50	159.50	0.2693	319	variant shiga-like toxin A subu
PIR2:176713	+	93.50	159.50	0.2693	319	variant shiga-like toxin A subu

seq\_name: pIR2:PD0018  
seq\_documentation\_block:  
mistletoe lectin I A chain - Viscum album (fragment)  
C:Species: Viscum album  
C:Date: 03-Jul-1998 #sequence\_revision 10-Jul-1998 #text\_change 26-Aug-1999  
C:Accession: PD0018  
R:Schubert, S.; Krauspenhaar, R.; Mikhailov, A.; Stoeva, S.; Betzel, C.; Voelker, W  
Biochem. Biophys. Res. Commun. 247, 367-372, 1998  
A:Title: Primary structure and molecular modelling of mistletoe lectin I from Viscum a  
A:Reference number: PD0018; MUID:98308123  
A:Accession: PD0018  
A:Molecule type: protein  
A:Residues: 1-254 <ESC>  
C:Superfamily: ricin  
F:7-246/Domain: RNA N-glycosidase homology <RNG>

alignment\_scores:  
Quality: 621.00 Length: 140  
Ratio: 4.634 Gaps: 0  
Percent Similarity: 95.714 Percent Identity: 85.714

alignment\_block:  
US-09-627-165b-15 x PD0018

Align seg 1/1 to: PD0018 from: 1 to: 254

1	TACACAGATCTGAGAGATACGCGGTCATAGGACACGATCCCTGGG	50
115	TYRPRASPLGGLNARGYRGLAGLYNARGASRGINLEPROLEUGL	131
51	TATAGAGAACATCATTCATCGCGGCGCTGTTATCCAGGCGCA	100
131	YLEASGGLNLEGLNLEGLNLEGLNLEGLNLEGLNLEGLNLEGLN	148
101	GCACCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	150
148	ERTHARGTHGALNARGSERLEULEULEULEULEULEULEULEULE	164
151	GAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	200
165	GLUALALALARGPHEASNPRIOLEUTPARGYRARGINTYRILEAS	181
201	CAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	250
181	NSEGLVALASERPHLEUPROASPVLYRMETLEUGLLEUGLTHRS	198
251	GTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	300
198	ERTPGGLGGLNLEGLNLEGLNLEGLNLEGLNLEGLNLEGLNLEGLN	214
301	AATACCATTCGTTGGGATATCCAGCGGCGGCGGCGGCGGCGG	350
215	ASNPASNPRIOLEUTGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY	231
351	CAATGTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	400
231	RANVALARGASPVLYLEALASERLEUALALILEMETLEUPHEVALC	248
401	GGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	420
248	LYGLNARGPERSERSESR	254

seq\_name: pIR1:TLZLSA  
seq\_documentation\_block:

abrin-a precursor - Indian licorice (fragment)  
N:Contains: RNA N-glycosidase (EC 3.2.2.22)  
C:Species: Abrus precatorius (Indian licorice)  
C:Date: 31-Dec-1993 #sequence, revision 01-Aug-1997 #text-change 16-Jul-1999  
C:Accession: S32429, J02020; A39761, J01398; S14472; S24133; S74110; S74111  
R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
J. Mol. Biol. 229, 263-267, 1993  
A:Title: Primary structure of three distinct isoabrisins determined by cDNA sequencing.  
A:Reference number: S32429; MUID:93132798  
A:Accession: S32429  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 'E', 2-528 <HUN>  
A:Cross-references: GB:M8344; NID:g166294; PIDN:AAA32624.1; PID:g166295  
A:Note: the coding region for the sequence shown is preceded by an ATG codon  
A:Note: residues 1-8 were derived from the synthesized primer  
R:Funatsu, G.; Taguchi, Y.; Kamenosono, M.; Yanaka, M.  
Agric. Biol. Chem. 52, 1095-1097, 1988  
A:Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein from  
A:Reference number: J02020  
A:Accession: J02020  
A:Molecule type: protein  
A:Residues: 1-201,203-251 <FUN>  
A:Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have  
R:Evensen, G.; Mathiesen, A.; Sundan, A.  
J. Biol. Chem. 266, 6848-6852, 1991  
A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.  
A:Reference number: A39761; MUID:91201328  
A:Accession: A39761  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 'E', 2-251 <EVE>  
A:Cross-references: GB:X54872  
R:Kimura, M.; Sumizawa, T.; Funatsu, G.  
Biosci. Biotechnol. Biochem. 57, 166-169, 1993  
A:Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic  
A:Reference number: J01398; MUID:93169023  
A:Accession: J01398  
A:Molecule type: protein  
A:Residues: 261-347, 'T', 349-351, 'A', 353-357, 'L', 359-528 <KIM>  
A:Experimental source: seed  
R:Evensen, G.; Mathiesen, A.; Sundan, A.  
submitted to the EMBL data Library, October 1990  
A:Description: Direct molecular cloning of two distinct abrin A-chains.  
A:Reference number: S14471  
A:Accession: S14471  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 'ME', 2-251 <EV2>  
A:Cross-references: EMBL:X54873; NID:g16090; PIDN:CAA38655.1; PID:g16091  
R:Chen, Y.L.; Chow, L.P.; Tsugita, A.; Lin, J.Y.  
FEBS Lett. 309, 115-118, 1992  
A:Title: The complete primary structure of abrin-a B chain.  
A:Reference number: S24133; MUID:92371656  
A:Accession: S24133  
A:Molecule type: protein  
A:Residues: 262-297, 'Y', 299-426, 'L', 428-466, 'P', 468-482, 'L', 484-528 <CHE>  
R:Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.  
Eur. J. Biochem. 240, 564-569, 1996  
A:Title: Probing the domain structure of abrin-a by tryptic digestion.  
A:Reference number: S74110; MUID:97008945  
A:Accession: S74110  
A:Molecule type: protein  
A:Residues: 89-108,154-172 <LIN>  
A:Experimental source: seed  
A:Accession: S74111  
A:Molecule type: protein  
A:Residues: 262-276, 'X', 278-280,329-348,369-388,399-418 <LIM>  
A:Experimental source: seed  
C:Comment: Abrin-a is more toxic than ricin. The toxin consists of an A chain, which in-  
taining receptors on the cell surface. The A and B chains are linked by a single disulfide

C:Superfamily: ricin; RNA N-glycoprotein; glycosidase; hydrolase; lectin; pyroglyutamic  
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglyutamic  
F:1-251/Product: abrin-a chain A #status experimental <ACH>  
F:7-246/Domain: abrin-a chain B #status experimental <BCB>  
F:261-328/Product: abrin-a chain B #status experimental <RCH>  
F:283-325,326-366,369-407,414-449,453-492,495-528/Region: 40-residue repeats  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
F:164,167/Active site: Glu, Arg #status predicted  
F:247-269,286-305,329-346,417-430,456-473/Disulfide bonds: #status predicted  
F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
F:361,401/Binding site: carboxylate (Asp) (covalent) #status experimental  
F:500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

## alignment\_scores:

Quality:	283.00	Length:	185
Ratio:	2.320	Gaps:	6
Percent Similarity:	65.946	Percent Identity:	38.919

## alignment\_block:

US-09-627-165b-15 x TZLSA ..

Align seg 1/1 to: TZLSA from: 1 to: 528

```
1  TACACGATCTGGACCATACGCGGTCTAT...AGGACGATCCCTCT 47
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
113  IYRGlyAspLeuGlnArgTrpAlaHisGlnSerArgGlnGlnIleProle 129
48  GGGTATAGAGCAATCATTCATTCGCGCTTCGCGCTTCGTTATCCAGCG 97
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129  uGlyLeuGlnAlaLeuThrHisGlyIleSerPhePheArgSerGlyGlyA 146
98  GCAGACCCGCGCCCAAGCTCGTCCCTTATATCTCATTCAGATGAGC 147
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146  snAspAsnGlnGlnIleValArgThrLeuIleValIleIleGlnMetVal 162
148  TCCGAGCGCGGAGATTCATTCATTCGCGCTTCGCGCTTCGCAATACAT 197
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
163  AlaglnAlaAlaArgPheAlaGlyTrpIleSerAsnArgValAlaValSerI 179
198  TAAACGCGGAGATTCATTCCTCCGACATGTACATGTACATGTACAGATG 247
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
179  eGlnThrGlyThrAlaPheGlnProAspAlaAlaMetIleSerLeuGlnA 196
248  CTAGTGGGCGCAACAATCCAGCAGTCCAGCAGTTCAGATGCGCTT 297
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196  snAsnTrpAspAsnLeuSerArgGlyValGlnIleSerValGlnAspThr 212
298  TTTAATATACCATTCGTTGGGTGATATCCACCGTAACCTTGTCAGCTT 347
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213  PheProAsnGln.....ValThrLe 219
348  GACCAATGTCGCGAC.....G 364
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
219  uThrAsnIleAlaGlnGlnProValIleValAspSerLeuSerHisProT 236
365  TGAATGCCAGCTTAGCATGATGTTGTATGTAGGACGACGACCATCT 414
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
236  hValAlaValIleValAlaMetLeuPheValGlyS...AsnProProAsn 251
415  TCCCTGACGATGCGCTATTCGCGCTGCTTCATTCAGACCGCTTGGAAAA 464
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
252  AlaAsnGlnSer.....ProLeuLeuIleArgSerIleValGlnIly 265
465  TAGCGGCGCGCTGACGATGTACCTGACGATGCTTCGCAACCGACCTGC 514
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
265  sSer.....LysIleCysSerSerArgTrpGlnProThrValA 278
515  GCATC 519
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
278  rgIle 279
```

seq\_name: p1r1:RLCSD

seq\_documentation\_block:

ricin D precursor-castor bean

N:Contains: Ricin N-glycosidase (EC 3.2.2.22)

C:Species: Ricinus communis (castor bean)

C:Date: 31-Dec-1993 #sequence-revision 31-Dec-1993 #text-change 16-Jul-1999

C:Accession: A24041; S20513; A24614; A03374; A24010; A03374; S10903

R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Iadin, B.F.; Houston, L.L.; Weaver, R.F.

C:Accession: A24041; S20513; A24614; A03374; A24010; A03374; S10903

Nucleic Acids Res. 13, 8019-8033, 1985

A:Title: Genomic cloning and characterization of a ricin gene from Ricinus communis.

A:Reference number: A24041; MUID:86067214

A:Accession: A24041

A:Molecule type: DNA

A:Residues: 1-576 <HML>

A:Cross-references: GB:X03179; MID:g21082; PIDN:CAA26939.1; PID:g21083

R:Tregear, J.W.; Roberts, L.M.

Plant Mol. Biol. 18, 515-525, 1992

A:Title: The lectin gene family of Ricinus communis: cloning of a functional ricin gene

A:Reference number: S20513; MUID:92163016

A:Accession: S20513

A:Molecule type: DNA

A:Residues: 1-576 <TRE>

A:Cross-references: EMBL:X52908; MID:g21084; PIDN:CAA37095.1; PID:g21085

R:Iamb, F.I.; Roberts, L.M.; Lord, J.M.

Eur. J. Biochem. 148, 265-270, 1985

A:Title: Nucleotide sequence of cloned cDNA coding for preproricin.

A:Reference number: A24614; MUID:85179479

A:Accession: A24614

A:Molecule type: mRNA

A:Residues: 12-75; 'D', 77-550; 'R', 552-576 <LAM>

A:Cross-references: GB:X03388; MID:g21077; PIDN:CAA26230.1; PID:g21078

R:Coshtake, S.; Funatsu, G.; Funatsu, M.

Agric. Biol. Chem. 42, 1267-1274, 1978

A:Title: Isolation and sequences of peptic peptides, and the complete sequence of the cDNA

A:Reference number: A03372

A:Accession: A03372

A:Molecule type: protein

A:Residues: 36-97; 'Q', 99-109; 'S', 111-269; 'D', 272-283; 'L', 285-288; 290-302 <YOS>

A:Note: this paper cites the others in the series providing experimental details for the

R:Nak, T.; Funatsu, G.

FEBS Lett. 191, 121-124, 1985

A:Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptophan

A:Reference number: A24010

A:Accession: A24010

A:Molecule type: protein

A:Residues: 315-383; 'PS', 386-576 <ARA>

R:Funatsu, G.; Kimura, M.; Funatsu, M.

Agric. Biol. Chem. 43, 2221-2224, 1979

A:Title: Primary structure of A chain of ricin D.

A:Reference number: A03374

A:Accession: A03374

A:Molecule type: protein

A:Residues: 315-335; 'N', 337-342; 'NH', 345-362; 364-383; 'PS', 386-399; 'T', 401; 'D', 403; 'E', 405; 'E', 529-564; 'W', 566; 'H', 567-570; 'L', 573-574; 'F', <FUN>

A:Note: this paper, one of a series, summarizes the experimental details for the determi

R:Reedy, M.P.; Kim, Y.; Robertus, J.D.

Protein 10, 270-278, 1991

A:Title: Site-directed mutagenesis of ricin A-chain and implications for the mechanism o

A:Reference number: A48237; MUID:91352006

A:Accession: A48237

A:Contents: annotation; active site

R:Rutenber, E.; Robertus, J.D.

Protein 10, 260-269, 1991

A:Title: Structure of ricin B-chain at 2.5 angstrom resolution.

A:Reference number: A48238; MUID:91352005

A:Accession: A48238

A:Contents: annotation; x-ray crystallography, 2.5 angstroms

R:Katzin, B.; Collins, E.; Robertus, J.D.

Protein 10, 251-259, 1991

A:Title: Structure of ricin A-chain at 2.5 angstroms.

A:Reference number: A48239; MUID:91352004

A:Accession: A48239

A:Contents: annotation; x-ray crystallography, 2.5 angstroms

C:Comment: The functional molecule is a disulfide-linked dimer of A and B chains, which

C:Comment: The A chain inhibits protein synthesis; it inactivates the 60S ribosomal subu

into the cell of the A chain; B chains are also responsible for cell agglutination (le  
C:Comment: This protein is cytotoxic and very poisonous to animals.  
C:Superfamily: ricin: RNA N-glycosidase homology  
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; s  
F:1-35/Domain: signal sequence #status predicted <SIG>  
F:36-300/Product: ricin D chain A #status experimental <ACH>  
F:46-293/Domain: RNA N-glycosidase homology <RNG>  
F:315-576/Product: ricin D chain B #status experimental <BCH>  
F:331-373; 374-414; 417-455; 462-497; 501-540; 543-576/Region: 40-residue repeats  
F:45; 409; 449/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:115; 158; 243; 244/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
F:212/Active site: Glu #status experimental  
F:215/Active site: Arg #status predicted  
F:294; 318; 334-353; 377-394; 465-478; 504-521/Disulfide bonds: #status experimental  
F:336; 349; 360/Binding site: N-acetylgalactosamine (Asp, Glu, Asn) #status experimental  
F:548; 569/Binding site: N-acetylgalactosamine (Asp, Asn) #status experimental

## alignment\_scores:

Quality:	259.50	Length:	178
Ratio:	2.257	Gaps:	5
Percent Similarity:	64.607	Percent Identity:	41.011

## alignment\_block:

US-09-627-165b-15 x RLCSD ..

Align seg 1/1 to: RLCSD from: 1 to: 576

```
1 TACACAGATCTGGAGGATACGCGCTCAT...AGGACCAAGATCCCTCT 47
|||||  |||||||  |||||||  |||||||  ||
158 TYRSPARGLEAGTUGLNUGLNLAALSERALAEUARYTYRYSER 174
48 GGGATATAGGAGATCTATATCCCTTCGCGCTCTCTT 90
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
174 UGLYASMLYPROLEUUGLNUGLNLAALSERALAEUARYTYRYSER 191
91 ..CCAGGGGAGCAGCCGCGGCGCTTCCTTATCTCAT 138
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
174 UGLYASMLYPROLEUUGLNUGLNLAALSERALAEUARYTYRYSER 191
191 HGLYGLYTHGLNLEUPROTHREULARGSERHELLEICYS 207
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
139 CAGATGATCTCGAGCGGAGATTCATCCATCTTTGGGCGCTG 188
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
208 GIMETILSERGLUALAALARGHELMYRIEGLNUGLNUET 224
189 CCATATCATATACCGGGAGATCTTCTCCGACATGATCATG 238
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
224 GTHARGLEARGTYRASHARGYSERLALAPROSPROSERVAL 241
239 AGCTGAGAGCTAGTGGGCGCACATCCAGCAAGTCACGACTAC 288
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
241 HLEUGLUNASERTPRGLYARGLSERTHRLAALIEGLNGLUS 257
289 GATGCGGTTTAAATACCATTCGTTGGTATATCCACCGTACT 338
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
258 GINGLYALAPHELASERPROLEUUGLNUGLNARGASNGLY 274
339 CGTACGCTTGACGATGTCGCGAGTGTATCCAGCTTACGAT 388
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
274 SPHESEVALTYRASPVALSERLLEUULEPROLEUULEUW 291
389 TGTGTATGATAGGACGACCATCTCTCCGACGTCGCTATTG 438
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
291 ALTYRARGYSALAPROPROSPROSERGLN.....PHESE 303
439 CTGTATATAGACCCGCTGTGAAATATACGCGCGCTGACGAT 488
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
304 LEULEULEARGPROVALVALPROASNPHEASNALA..... 317
489 CTGACCTCTTCGACCAACCGCGGCGCATCGTA 522
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
318 CYSEWASPPOGLNUGLNUGLNLAALIEGLNGL 328
```

seq\_name: p1r2:S32431

## seq\_documentation\_block:

abrin-d precursor - Indian licorice (fragment)  
 M/contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C/Species: Abrus precatorius (Indian licorice)  
 C/Date: 30-Sep-1993 #sequence\_revision 01-Aug-1997 #text\_change 01-Aug-1997  
 C/Accession: S32431; S34408  
 R/Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
 J. Mol. Biol. 229, 263-267, 1993  
 A>Title: Primary structure of three distinct isoabtrins determined by cDNA sequencing. C  
 A/Reference number: S32429; MUID:93132798  
 A/Accession: S32431  
 A/Molecule type: mRNA  
 A/Residues: 1-528 <HUN>  
 A/Cross-references: GB:M98346  
 R/Hung, C.; Lee, M.; Lee, T.; Lin, J.  
 submitted to the EMBL Data Library, March 1993  
 A/Reference number: S34408  
 A/Accession: S34408  
 A/Molecule type: mRNA  
 A/Residues: 1-169, 'C', 171-320, 'L', 322-528 <H02>  
 A/Cross-references: GB:M98346  
 C/Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating  
 the A and B chains are linked by a single disulfide bond, which is essential for toxic  
 C/Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; F  
 F/1-231/Product: abrin-d chain A #status predicted <ACH>  
 F/17-246/Domain: rRNA N-glycosidase homology <RNG>  
 F/261-528/Product: abrin-d chain B #status predicted <BCH>  
 F/283-325, 326-366, 369-407, 414-449, 453-492, 495-528/Region: 40-residue repeats  
 F/1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted  
 F/174, 113, 195, 196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F/164, 167/Active site: Glu, Arg #status predicted  
 F/200, 253, 361, 401, 402/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F/247-269, 286-305, 329-346, 417-430, 456-473/Disulfide bonds: #status predicted  
 F/288, 312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
 F/500, 521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

## alignment\_scores:

Quality: 256.00 Length: 175  
 Ratio: 2.151 Gaps: 5  
 Percent Similarity: 68.000 Percent Identity: 40.000

alignment\_block:  
 US-09-627-165B-15 x S32431 ..

Align seg 1/1 to: S32431 from: 1 to: 528

```

1  TACACGATCTGGAGCGATACGCCGGTCAT..AGGACCAAGATCCCTCT 47
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113  TGTGlyAspLeuGlnAlaArgTyrPheValHisGlnThrArgGlnGluIleSerIle 129
48  GGGTATAGAGAACTCATTCATCCGCTCTCGCGCTTCGTTATCCAGGCG 97
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
129  uGlyLeuGlnAlaLeuThrHisAlaIleSerPheLeuAlaGlySerGlyAla 146
98  GCAGCACCCGGGCCAAGCTGTCGCTTATATCTCATTCAGATGATC 147
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
146  erAsnAspGlnGluIleValAlaArgThrLeuIleValIleIleGlnMetAla 162
148  TCCGAGCGCGGAGATTCATCCCATCTTTGGAGGGCTCGCCAAATACAT 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163  SerGlnAlaAlaAlaArgTyrAlaTyrIleSerAsnArgValAlaIleSerIle 179
198  TAACAGCGGGAGATCTTCTCCGACATGTACATGTCGAGTCGAGAGA 247
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179  eArgThrGlyThrAlaPheGlnProAspProAlaMetLeuSerLeuGlnAla 196
248  CTGAGTGGGGCCAAATCCAGCAATCCAGCAATCCAGCAATCCAGTGCCTT 297
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196  snAsnThrPaspAsnLeuSerIleGlyValGlnIleGlnIleSerValGlnAspThr 212

```

```

298  TTT...AATAACCAATTTCGTTGGGTATATACACGGGTAATCTCGTAC 344
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
213  PheProAsnValIleLeuSerSerIleAsnArgGlnProValValIleVal 229
345  GTTGACCAATGTTCGCGAGTATCGCACGCTTACCATCATGTTGTTTG 394
   : ||| : : : : : ||||| ||||| ||||| ||||| ||||| |||||
229  LaspSerLeuSerHisProThrValAlaValIleValLeuMetLeuPheVal 246
395  TATGATAGGAGACCGACATCTTCCCTCCAGCTGCGCTATTTGGCCGCTGATC 444
   ||||| : : : ||||| ||||| ||||| ||||| ||||| ||||| |||||
246  AlGys...AsnProProAsnAlaAsnGlnSer.....ProLeuLeu 258
445  ATACGACCCGCTTGGAAATAGCGGGCGCGCTGACGATGTATACCTGCAC 494
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
259  IleArgSerIleValGlnGluSer.....LysIleCysSer 271
495  TGCTTCCGACCCACCGCTGCGCATC 519
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
271  rArgTyrGlnProThrValArgIle 279

```

seq\_name: p1r2:S16022

## seq\_documentation\_block:

abrin-c precursor - Indian licorice  
 M/contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C/Species: Abrus precatorius (Indian licorice)  
 C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Aug-1999  
 C/Accession: S16022  
 R/Wood, K.A.; Lord, J.M.; Mawczynczak, E.J.; Platak, M.  
 Eur. J. Biochem. 198, 723-732, 1991  
 A>Title: Preproabrin: genomic cloning, characterisation and the expression of the A-c  
 A/Reference number: S16022; MUID:91266957  
 A/Accession: S16022  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-562 <RNO>  
 A/Cross-references: EMBL:X55667; NID:q16084; PIDN:CA39202.1; PID:q16085  
 C/Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating  
 the A and B chains are linked by a single disulfide bond, which is essential for fox  
 C/Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglutamic a  
 F/35-285/Product: abrin-c chain A #status predicted <ACH>  
 F/41-280/Domain: rRNA N-glycosidase homology <RNG>  
 F/295-562/Product: abrin-c chain B #status predicted <BCH>  
 F/317-359, 360-400, 403-441, 448-483, 487-526, 529-562/Region: 40-residue repeats  
 F/35/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
 F/108, 147, 229, 230/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F/198, 201/Active site: Glu, Arg #status predicted  
 F/234, 287, 395, 435, 436/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F/281, 303, 320-339, 365-380, 451-464, 490-507/Disulfide bonds: #status predicted  
 F/322, 346/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
 F/534, 555/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

## alignment\_scores:

Quality: 256.00 Length: 175  
 Ratio: 2.151 Gaps: 5  
 Percent Similarity: 68.000 Percent Identity: 40.000

alignment\_block:  
 US-09-627-165B-15 x S16022 ..

Align seg 1/1 to: S16022 from: 1 to: 562

```

1  TACACGATCTGGAGCGATACGCCGGTCAT..AGGACCAAGATCCCTCT 47
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
147  TGTGlyAspLeuGlnAlaArgTyrPheValHisGlnThrArgGlnGluIleSerIle 163
48  GGGTATAGAGAACTCATTCATCCGCTCTCGCGCTTCGTTATCCAGGCG 97
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163  uGlyLeuGlnAlaLeuThrHisAlaIleSerPheLeuAlaGlySerGlyAla 180
98  GCAGCACCCGGGCCAAGCTGTCGCTTATATCTCATTCAGATGATC 147

```







421 .....GACGTGGCGTATGCGCGGTGTCATACGACCCGCTTGG 461  
||||| : : : : :  
295 AsnAlaIleSpAlaIn.....IleIleIleYMetProAlaIleA 308  
462 AAATAGCGCGCGCGTGCAGCATTTACCTGCTCCGACCAACCG 511  
: : : : :  
308 g.....GlyGlyGlyTyrGlySerValGlyValThr 323  
512 TCGCGATC 519  
|||||  
323 rGaArgIle 325  
seq\_name: p1r2:C39761

seq\_documentation\_block:  
abrin (clone 7.2) precursor - Indian licorice (fragment)  
N:contains: rRNA N-glycosidase (EC 3.2.2.22)  
C:Species: Abrus precatorius (Indian licorice)  
C:Date: 21-Feb-1992 #sequence\_revision 31-Dec-1993 #text\_change 20-Aug-1999  
C:Accession: C39761; S14471  
R:Evensen, G.; Mathiesen, A.; Sundan, A.  
J. Biol. Chem. 266, 6848-6852, 1991  
A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.  
A:Reference number: A39761; MUID:91201329  
A:Accession: C39761  
A:Molecule type: DNA  
A:Residues: 1-251 <EVE>  
R:Evensen, G.; Mathiesen, A.; Sundan, A.  
submitted to the EMBL Data Library, October 1990  
A:Description: Direct molecular cloning of two distinct abrin A-chains.  
A:Reference number: S14471  
A:Accession: S14471  
A:Molecule type: DNA  
A:Residues: 'M', 1-251 <EVE>  
A:Cross-references: EMBL:X54872; NID:g16088; PIDN:CAA3854.1; PID:g16089  
C:Superfamily: ricin; rRNA N-glycosidase homology  
C:Keywords: duplication; glycosidase; hydrolase; lectin; toxin  
F:1-251/Product: abrin (clone 7.2) chain A #status predicted <ACH>  
F:7-246/Domain: rRNA N-glycosidase homology <RNG>  
F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
F:164,167/Active site: Glu, Arg #status predicted

alignment\_scores:  
Quality: 220.00 Length: 135  
Ratio: 2.316 Gaps: 2  
Percent Similarity: 70.370 Percent Identity: 40.741

alignment\_block:  
US-09-627-165B-15 x C39761 ..

Align seg 1/1 to: C39761 from: 1 to: 251

1 TACACAGATCTGAGCGATACGCGGTCAT...AGGACCAATCCCTCT 47  
||| ||||| : : : : :  
113 TysGlyAspLeuGluArgTyrPAlaIleSerPheLeuArgSerIle 129  
GGGTATAGGAGCACTTCAATCGGCGCTGCGGCTGATATCCAGCGC 97  
||||| : : : : :  
129 uGlyLeuAlaLeuThrHisAlaIleSerPheLeuArgSerIle 146  
GCAGACACCGGCGCCAGAGCTGTTCCCTTATATCTCATTCAGATGTC 147  
: : : : :  
146 eAsnAspGluGlnLysAlaArgThrLeuIleValIleIleGlnMetAla 162  
148 TCCAGAGCGCGAGATTCATCCATCTTTGGAGGCGTCCCAATACAT 197  
||||| : : : : :  
163 SerGlnAlaAlaArgTyrArgTyrIleSerAsnArgValGlyValSer 179  
198 TACAGCGGCGGAGATCTTCCCGACATGATGATGATGATGATGATG 247  
| : : : : :  
179 eArgThrGlyThrAlaPheGlnProAspProAlaMetLeuSerLeuGln 196

248 CTAGTTGGGCGCCAGACATCCAGCAAGTCCAGACACTACGATGCGCTT 297  
: : : : :  
196 snAsnTrpAspAsnLeuSerArgGlyValGlnGlnIleSerValGlnAsp 212  
298 TTTTAAATACCCATTTTCGTTG...CGTATATCCACCGGTAATTGCTAC 344  
||| ||| : : : : :  
213 PheProAsnAlaValThrLeuArgArgValAsnAsnGlnProValIleVal 229  
345 GTTGGACCAATGTTGGGAGCTGATCCAGCTTACGATGATGATGATG 394  
: ||| : : : : :  
229 LaspSerLeuThrHisGlnSerValAlaValLeuAlaLeuMetLeuPhe 246  
395 TATGT 399  
|||  
246 aLCys 247

seq\_name: p1r2:B39761  
seq\_documentation\_block:  
abrin (clone 3.7) precursor - Indian licorice (fragment)  
N:contains: rRNA N-glycosidase (EC 3.2.2.22)  
C:Species: Abrus precatorius (Indian licorice)  
C:Date: 21-Feb-1992 #sequence\_revision 31-Dec-1993 #text\_change 05-Jun-1998  
C:Accession: B39761  
R:Evensen, G.; Mathiesen, A.; Sundan, A.  
J. Biol. Chem. 266, 6848-6852, 1991  
A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.  
A:Reference number: A39761; MUID:91201329  
A:Accession: B39761  
A:Molecule type: DNA  
A:Residues: 1-106 <EVE>  
A:Cross-references: GB:X54873  
C:Superfamily: ricin; rRNA N-glycosidase homology  
C:Keywords: duplication; glycosidase; hydrolase; lectin; toxin  
F:1-106/Product: abrin chain A (fragment) #status predicted <ACH>  
F:1-101/Domain: rRNA N-glycosidase homology (fragment) <RNG>  
F:19,22/Active site: Glu, Arg #status predicted

alignment\_scores:  
Quality: 180.50 Length: 107  
Ratio: 2.654 Gaps: 2  
Percent Similarity: 63.551 Percent Identity: 41.121

alignment\_block:  
US-09-627-165B-15 x B39761 ..

Align seg 1/1 to: B39761 from: 1 to: 106

112 CAAGCTCGTTCCTTATATCCATTCATTCAGATGATCTCGAGCGCGAG 161  
: : : : :  
6 LysAlaArgThrLeuIleValIleIleGlnMetValSerGlnAlaIle 22  
162 ATTCATCCATCTTTGGAGGCGTCCCAATATCAATACAGCGGAGAT 211  
||| : : : : :  
22 gPheAlaGlyIleSerAsnArgValGlyValSerIleArgThrGlyThr 39  
212 CATTTCTCCGACATGATGATCTGAGCTGAGCTGAGCTGAGCTGAG 261  
: ||| : : : : :  
39 LaspGlnProAspProAlaMetIleSerLeuGlnAsnTrpAspAsn 55  
262 CAATCCAGCAAGTCCAGATGATGATGATGATGATGATGATGATG 311  
||| : : : : :  
56 LeuSerGlyGlyValGlnGlnSerValGlnGlyThrPheProHisPro 71  
312 TCGGTTGGTATATCCACCGGTAATCTCGAGAGTTCAGCAATGTT 357  
||| : : : : :  
72 .....ValThrLeuArgAsnIleCysA 79  
358 .....CGCAGCTGATCGCCAGCTTA 378  
| : : : : :  
79 snGlnProValIleValAspSerLeuSerHisProThrValAlaValLeu 95

```

seq_name: p1r1:RLT2T
seq_documentation_block:
  RNA-N-glycosidase (EC 3.2.2.22) alpha-trichosanthin precursor [validated] - Mongolian snake-head
  N-Alternate names: alpha-TCs; type I ribosome-inactivating protein
  C:Species: Trichosanthes kirilowii (Mongolian snake-gourd)
  C:Date: 30-Sep-1988 #sequenceRevision 26-Jan-1996 #textChange 23-Mar-2001
  C:Accession: J00566; A36274; J01093; A36273; J00003
  R:Shaw, P.C.; Yung, M.H.; Zhu, R.H.; Ho, W.K.K.; Ng, T.B.; Yeung, H.W.
  Gene 97, 267-272, 1991
  A:Title: Cloning of trichosanthin cDNA and its expression in Escherichia coli.
  A:Reference number: J00566; MUID:91153657
  A:Accession: J00566
  A:Molecule type: mRNA
  A:Residues: 1-289 <SHA>
  A:Cross-references: GB:M34858; NID:g170536; PIDN:AAA34207.1; PID:g170537
  A:Experimental source: tuber
  R:Chow, T.P.; Feldman, R.A.; Lovett, M.; Piatek, M.
  J. Biol. Chem. 265, 8670-8674, 1990
  A:Title: Isolation and DNA sequencing of a gene encoding alpha-trichosanthin, a type I ribosome-inactivating protein.
  A:Reference number: A36274; MUID:90256790
  A:Accession: A36274
  A:Molecule type: DNA
  A:Residues: 1-233,'T',235-246,'M',248-289 <CHO>
  A:Cross-references: GB:J05434; NID:g170534; PIDN:AAA34206.1; PID:g170535
  R:Zheng, H.G.; Wang, B.; Shao, P.Z.; Yang, X.R.
  Acta Genet. Sin. 21, 42-51, 1994
  A:Title: Cloning and DNA sequencing of the gene encoding Trichosanthin.
  A:Reference number: J01093; MUID:94271613
  A:Accession: J01093
  A:Molecule type: DNA
  A:Residues: 1-172,'V',74-90,'S',92-233,'T',235-267,'D',269-289 <ZHE>
  A:Cross-references: GB:S70176; NID:g547148; PIDN:AB31048.1; PID:g547149
  R:Collins, E.J.; Robertus, J.D.; Lofrest, M.; Stone, K.L.; Williams, K.R.; Wu, P.; Hwang, J.
  J. Biol. Chem. 265, 8665-8669, 1990
  A:Title: Primary amino acid sequence of alpha-trichosanthin and molecular models for alpha-trichosanthin.
  A:Reference number: A36273; MUID:90256789
  A:Accession: A36273
  A:Molecule type: protein
  A:Residues: 24-270 <COL>
  R:Wang, Y.; Qian, R.Q.; Gu, Z.W.; Jin, S.W.; Zhang, L.Q.; Xia, Z.X.; Tian, G.Y.; Ni, C.Z.
  Pure Appl. Chem. 58, 789-798, 1986
  A:Title: Scientific evaluation of Tian Hua Fen (THF): history, chemistry and application.
  A:Reference number: J00003
  A:Accession: J00003
  A:Molecule type: protein
  A:Residues: 24-56,'L',58-59,'I',61-71,'I',73-81,85-86,'L',88-92,'DAGLPNAV',93-142,'GL'
  A:Experimental source: tuber
  R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.
  submitted to the Brookhaven Protein Data Bank, July 1994
  A:Contents: annotation; X-ray crystallography, 1.6 angstroms, with adenine, residues 24-142
  R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.
  submitted to the Brookhaven Protein Data Bank, July 1994
  A:Reference number: A67092; PDB:1MKK
  A:Contents: annotation; X-ray crystallography, 1.6 angstroms, with formycin, residues 24-142
  R:Xiong, J.P.; Xia, Z.X.; Wang, Y.
  submitted to the Brookhaven Protein Data Bank, December 1994
  A:Reference number: A66711; PDB:1RCS
  A:Contents: annotation; X-ray crystallography, 1.7 angstroms, with NADPH, residues 24-27
  R:Xiong, J.P.; Xia, Z.X.; Wang, Y.
  Nat. Struct. Biol. 1, 695-700, 1994
  A:Title: Crystal structure of trichosanthin-NADPH complex at 1.7 Angstroms resolution.
  A:Reference number: A58622; MUID:95360714
  A:Contents: annotation; X-ray crystallography, 1.7 angstroms
  C:Comment: Alpha-trichosanthin has been used to induce abortions.
  C:Genetics:
  A:Gene: tcs
  A:Function:

```

[illegible]

Biol. Pharm. Bull. 20, 711-713, 1997  
A:Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating prote  
A:Reference number: JCS606; MUID:97356562  
A:Accession: JCS606  
A:Molecule type: DNA  
A:Residues: 1-289 <MIZ>  
A:Cross-references: DDBJ:AB000666; NID:92329830; PID:BA21786.1; PID:92329831  
R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.  
Biol. Pharm. Bull. 19, 1485-1489, 1996  
A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka  
A:Reference number: JCS032; MUID:97108848  
A:Accession: JCS033  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 22-270 <KON>  
C:Comment: This protein is a ribosome-inactivating protein and exhibit cytotoxic, aborti  
C:Comment: This protein belongs to type I ribosomal-inactivating proteins which catalyze  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: glycosidase  
F:22-270/Product: karasurin C #status predicted <MNC>  
F:24-270/Product: karasurin A #status predicted <MAA>  
F:27-266/Domain: rRNA N-glycosidase homology <RNG>

alignment\_scores:  
Quality: 157.00 Length: 164  
Ratio: 1.570 Gaps: 5  
Percent Similarity: 60.976 Percent Identity: 29.878

alignment\_block:  
US-09-627-165B-15 x JCS606 ..

Align seg 1/1 to: JCS606 from: 1 to: 289

```
1 TACACAGTCGTGAGACGATACGCGGTCTAT...AGGACGACGATCCGCT 47
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134 Tyrgluatrgleuglnlealaalaaglylellearglnuhenilleprole 150
GGGTATAGAGAGACATTCATCCGCTCGCGCTTCGTTATCCAGCGC 97
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
150 uGlyleuprolaleuaspseralallethrtleuphetrytrysna 167
GCAGCAGCCGCGGCGGAGCTCTCCCTTATATCATTCATTCAGATGATC 147
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 laaenSerAlaAlaSerAla.....leuMetValleuilegInserThr 181
TCCGAGCGCGGAGATTCATCCATCTTTGGAGCGCTGCCAATACAT 197
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 SerGluaAlaAlaArgTyrLysPheilegluglnlelleglylsarGva 198
198 TAAACAGCGGCGGAGTCATTTCTCCGACATGTACATGCTGAGCTGAGA 247
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
198 lAsp.....lysthrPheleuproSerleuAlaAlaileSerleuAla 213
248 CTATGTTGGGCGCAACAATCCAGCAGATCCAG.....CAGTTCAGCAT 291
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 snSerTTPSerAlaAlaSerleuSerleuilegInleAlaSerThrAsn 229
292 GCGCTTTTATATACCATTTCCGTTGGGTATATCCACCGTAACTTCGT 341
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
230 GlyInPheglutntrProvalValleuileasnaAlaInasnGlnarGva 246
342 GACGTTGAGCAATGTTCCGACGATGATGCCGCTTACGATCATGTTGT 391
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
246 lThrIleThrAsnValaspAlaIyValValThrSerAsnIleAlaLeuL 263
392 TTGTATGTAGGAGCGGACCATCTTCTCGACGTGCGTATTTGCGCGCTG 441
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
263 euLeuAnaArgasn..... 267
442 GTCATACGACCGCTCTTGAAATATAGCGCGCGTGAGCAT 483
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
268 .....AsnMetAlaAlaIleAsp 274
```

seq\_name: plr2:JE0401

seq\_documentation\_block:  
antiviral protein - Virginian pokeweed  
C:Species: Phytolacca americana (Virginian pokeweed)  
C:date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 07-May-1999  
C:Accession: JE0401  
R:Kung, S.S.; Kimura, M.; Funatsu, G.  
Agric. Biol. Chem. 54, 3301-3318, 1990  
A:Title: The complete amino acid sequence of antiviral protein from the seeds of poke  
A:Reference number: JE0401; MUID:91242096  
A:Accession: JE0401  
A:Molecule type: protein  
A:Residues: 1-261 <KUN>  
A:Experimental source: seed  
C:Comment: This protein prevents the replication of a number of plant viruses, and in  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: disulfide bond; glycoprotein  
F:6-254/Domain: rRNA N-glycosidase homology <RNG>  
F:10,44,255/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:34-258,84-105/Disulfide bonds: #status experimental

alignment\_scores:  
Quality: 156.50 Length: 135  
Ratio: 1.720 Gaps: 3  
Percent Similarity: 67.407 Percent Identity: 30.370

alignment\_block:  
US-09-627-165B-15 x JE0401 ..

Align seg 1/1 to: JE0401 from: 1 to: 261

```
1 TACACAGATCTGGAGGATATACGCGGT.....CATAGGACCAAGATCCC 44
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 TyrProThrleuGlnLysLysAlaGlyValThrSerArgasnIleVala 138
45 TCGGGTATAGAGAGACATTCATCCGCTCGCGCTTCGTTATCCAG 94
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 nleuglyleleGlnleleuSerSeraspIleleuSerleuValaIleGln 155
95 GCGGCGACGACC...CGGCGCAAGCTCGTCCCTTATATCTCATTCAG 141
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
155 lYserPheThrGlnLysIleGlnAlaLysPheleuValaIleGln 171
142 ATGATCTCCGAGCGCGGAGATTCATCCATCTTTGGAGGCGTCCGCA 191
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
172 MetValSerGlnAlaAlaArgPheLysTyrIleGlnAsnGlnValLys.. 187
192 ATACATTAACAGCGGAGTCATTTCTCCGACATGTACATGCTCGAGC 241
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
188 ....ThrAsnPheAsnArgAspPheSerProAsnAspLysValleuAspL 203
242 TGGAGACTGTTGGGCGCAACAATCCAGCAGATCCAGATCTTCAGCAT 291
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
203 euGlnLysasntrPglLysIleSerThrAlaIleHisAsnSerLysAsn 219
292 GCGCTTTTATATACCATTTCCGTTGGGTATATCCACCGTAACTTCGT 341
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
220 GlyAlaLeuProLysProleuGlnleuLysAsnAlaAspGlyThrLysTr 236
342 GACGTTGAGCAATGTTCCGACGATGATGCCGCTTACGATCATGTTGT 391
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
236 plleValleuArgValaspGlnleleLysProaspAlaGlyleuLeuAsnT 253
392 TTGTA 396
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
253 yTrVal 254
```

seq\_name: plr2:S28421  
seq\_documentation\_block:

rRNA N-glycosidase (EC 3.2.2.22) PAP-alpha - Virginian pokeweed  
 M:Alternate names: antiviral protein alpha PAP  
 C:Species: Phytolacca americana (Virginian pokeweed)  
 C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 20-Jun-2000  
 C:Accession: S28421  
 R:Katoka, J., Habuka, N.; Masuta, C.; Miyano, M.; Kotwal, A.  
 Plant Mol. Biol. 20, 879-886, 1992  
 A:Title: Isolation and analysis of a genomic clone encoding a pokeweed antiviral protein  
 A:Reference number: S28421; MUID:93099240  
 A:Accession: S28421  
 A:Molecule type: DNA  
 A:Residues: 1-294 <RAT>  
 A:Cross-references: EMBL:D10600; NID:9218010; PIDN:BA01451.1; PID:9218011  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 C:Keywords: glycosidase; hydrolase  
 F:30-278/Domain: rRNA N-glycosidase homology <RNG>

alignment\_scores:  
 Quality: 156.00 Length: 134  
 Ratio: 1.660 Gaps: 3  
 Percent Similarity: 70.149 Percent Identity: 30.597

alignment\_block:  
 US-09-627-165B-15 x S28421 ..

Align seg 1/1 to: S28421 from: 1 to: 294

```

1 TACACAGATCTGAGCGATACCGCGTCAT...AGGCACGATCCCTCT 47
||| |||||: :|||: |||||: |||
147 TyrProAlaLeuGluLysValGlyYrProAlrGserGlnValGlnLe 163
48 GGGTATAGAGAGACTCAATTCAGTCCGCTCGCGCTTCGTATCCAGGCG 97
||| |||||: :|||: |||||: |||
163 uGlyLeuGlnLeuAsnSerGlyLeuGlyLysLeuGlyValAspS 180
98 GCAGCACC...CGGCGCCAGCTGCTCCCTTATATCCATTCATTCAGATG 144
: ||| :|||: |||||: |||||: |||||
180 erPheThrGlnLysThrGlnLagLuphLeuLeuValAlaLeuGlnmet 196
145 ATCTCCGAGCGCGGAGATTCATCCATCCATCTTTTGAGAGGCTCCGCAATA 194
: |||||: |||||: |||||: |||||: |||||
197 ValSerGlnAlaAlaArgPheLysThrLleGlnAsnGlnValLys.... 211
195 CATTAACAGCGGGGAGATCATTTCTCCCGACATGTACATGCTCCAGAGTGG 244
||| :|||: |||||: |||||: |||||
212 .ThrAsnPhenAlaArgAlaPheTyTrProAsnAlaLysValLeuAsnLeuG 228
245 AGACTAGTTGGGGCCACATCCACGCAAGTCCAGCAGCTTACGAGATGGC 294
|| |||||: |||||: |||||: |||||: |||||
228 LuGlnSerTrpGlyLysLleSerThrAlaLeuAsnAlaLysAsnGly 244
295 GTTTTATTAACCCATTTCCGTTGGGTATATCCACCGGTAACCTCGTCAC 344
: |||||: |||||: |||||: |||||: |||||
245 AlaLeuThrSerProLeuGlnLeuLysAsnAlaLysGlySerLysTrpI 261
345 GTTGAGCAATGTTCCGAGCGATGAGCCAGCTTACGATCATGTGTTGTTGG 394
: ||| :|||: |||||: |||||: |||||
261 eValLeuArgValAspAspIleGluProAspValGlyLeuLeuLysTyTrV 278
395 TA 396
||
278 al 278

```

seq\_name: plr2:JC5032

seq\_documentation\_block:  
 karasurin-B - Trichosanthes kirilowii var. japonica  
 C:Species: Trichosanthes kirilowii var. japonica  
 C:Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 23-May-1997  
 C:Accession: JC5032  
 R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.  
 Biol. Pharm. Bull. 19, 1485-1489, 1996

A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and  
 A:Reference number: JC5032; MUID:97108848  
 A:Accession: JC5032  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-247 <KON>  
 C:Comment: This protein belongs to type I ribosomal-inactivating proteins which catal  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 F:4-243/Domain: rRNA N-glycosidase homology <RNG>

alignment\_scores:  
 Quality: 154.50 Length: 135  
 Ratio: 1.644 Gaps: 5  
 Percent Similarity: 69.630 Percent Identity: 33.333

alignment\_block:  
 US-09-627-165B-15 x JC5032 ..

Align seg 1/1 to: JC5032 from: 1 to: 247

```

1 TACACAGATCTGAGCGATACCGCGTCAT...AGGCACGATCCCTCT 47
||| |||||: |||||: |||||: |||||
111 TyrGluArgLeuGlnLeuAlaGlyLysLleArgGlnAsnLeuProle 127
48 GGGTATAGAGAGACTCAATTCAGTCCGCTCGCGCTTCGTATCCAGGCG 97
|||||: ||| :|||: |||||: |||||
127 uGlyLeuProAlaLeuAspSerAlaLleThrLeuPheTyTrYrAsnA 144
98 GCAGCACC...CGGCGCCAGCTGCTCCCTTATATCCATTCATTCAGATGATC 147
: |||||: |||||: |||||: |||||: |||||
144 laAsnSerAlaAlaSerAla....LeuMetValLeuLleGlnSerThr 158
148 TCCGAGGCGCGAGATTCATCCATCCATCTTTTGAGAGGCTCCGCAATACAT 197
||| |||||: |||||: |||||: |||||
159 SerGlnAlaAlaArgTyTrLysPheLleGlnGlnLleGlyLysArgTyA 175
198 TAACAGCGGGGAGTCTTCTCCCGACATGTACATGCTCGAGCTGGAGA 247
: |||||: |||||: |||||: |||||: |||||
175 laSp.....LysThrPheLeuProSerLeuAlaLleLleSerLeuGlnA 190
248 CTAGTGGGGCCACATCCAGCAGATCCAG....CAGTCTACGAT 291
: |||||: |||||: |||||: |||||: |||||
190 snSerTrpSerAlaLeuSerLysGlnLleGlnLleAlaSerTyTrAsnAsn 206
292 GCGGTTTATTAACCCATTTCCGTTGGGTATATCCACCGGTAACCTTCGT 341
||| |||||: |||||: |||||: |||||
207 GlyGlnPheGluThrProValValLeuLleAsnAlaGlnAsnLargTyA 223
342 GACGTTGAGCAATGTTCGCGAC.....GTGATGCCAGCTTAGCGATCA 385
|||||: |||||: |||||: |||||: |||||
223 lThrLleThrAsnValAlaSnAlaGlyValAlaThrSerAsnLleAlaLeuL 240
386 TGTGG 390
: |||
240 euLeu 241

```

**us-09-627-165b-15.rsp**

SwissProt  
SwissProtSwissPro  
GuttedProt

SwissPro  
sec name

seq_documentation_block:				
ID	MLA_VISAL	STANDARD:	PRT;	254 AA.
AC	P81446:			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Beta-galactoside specific lectin I A chain (MLA) (ML-I A) (rRNA N-glycosidase) (EC 3.2.2.22).			
OS	Viscum album (European mistletoe).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Santalales; Viscaceae; Viscum.			
CC	Went 1949D-2012			

RN (1)  
 RP SEQUENCE.  
 RC STRAIN-SUBSP. ALBUM;  
 RX MEDLINE-97134581; PubMed-8980141;  
 RA Hugert Soler M., Stoeva S., Schamborn C., Wilhelm S., Stiefel T.,  
 RT Voelter W.;  
 \*Complete amino acid sequence of the A chain of maitotoxin lectin I.\*;  
 3600153-15721965

CC SUBUNIT'S BY REMOVING ADENINE FROM POSITION 4, 5, 24 OF 20 3 NUCLEOTIDE AND  
CC  
CC P CHAIN BINDS TO CELL RECEPTORS AND PROBABLY FACILITATES THE ENTRY

```

CC      B CHAIN BRINGS TO CELL AGGLUTINATION. B CHAINS ARE ALSO RESPONSIBLE FOR
CC      INTO THE CELL OF THE A CHAIN; B CHAINS ARE ALSO RESPONSIBLE FOR
CC      CELL AGGLUTINATION (LECTIN ACTIVITY).
CC      CARLYATIC ACTIVITY: Endoglycolysis of the N-glycosidic bond at one
CC      specific asenosine on the 28S rRNA.
CC      -1 SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC      -1 MISCELLANEOUS: TWO ISOPFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA
CC      AND NON-GLYCOSYLATED FORM MLA.
CC      -1 SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 2
CC      RIP SUBFAMILY.
CC      HSP: P11140; IABR.
DR      Interpro: IPR001574; RIP.
DR      Pfam: PF00161; RIP; 1.
DR      PRINTS: PR00396; SHIGARICIN.
DR      PROSITE: PS00275; SHIGA_RICIN; FALSE_NEG.
KW      Hydrolase; Toxin; Repeat; Glycoprotein; Lectin.
FT      ACT_SITE 165
FT      CARBOHYD 112 112
FT      VARIANT 15 15
FT      VARIANT 66 66
FT      VARIANT 112 112
FT      VARIANT 116 116
FT      VARIANT 133 134
FT      VARIANT 140 140
FT      VARIANT 144 144
FT      VARIANT 151 151
FT      VARIANT 179 179
FT      VARIANT 184 184
FT      VARIANT 190 190
FT      VARIANT 218 218
FT      VARIANT 223 224
FT      VARIANT 231 231
FT      VARIANT 235 235
SQ      SEQUENCE 254 AA; 28478 MW; 53BAF98D3E0FEE67 CRC64;
alignment_scores:
Quality: 621.00
Ratio: 4.634
Percent Similarity: 95.714
Length: 140
Gaps: 0
Percent Identity: 85.714

```



## alignment\_block:

US-09-627-165b-15 x MLA\_VISAL ..

Align seg 1/1 to: MLA\_VISAL from: 1 to: 254

```

1 TACACGATCTGGAGCGATACCGCGTCATAGGACGACGATCCCTGGG 50
115 TYPProAspLeuGluArgTyrAlaGlyHisArgAspGlnIleProLeuG 131
51 TATAGAGAACTCATTCATCCGTCGGCGCTTCGTCATACGACGGCA 100
131 YLLeAspGlnLeuIleGlnSerValThrAlaLeuArgPheProGlyG 148
101 GCACCGGGCCCAAGCTCGTCCCTTATATACCTCATTCAGATGCTCC 150
148 eTThrArgThrGlnAlaArgSerIleLeuIleGlnMetIleSer 164
151 GAGGGCGGAGATTCATCCATCTTTGGAGCGCTGCCCATCATTTAA 200
165 GluAlaAlaArgPheAsnProIleLeuThrArgTyrArgGlnTyr 181
201 CAGCGGGAGTCATTTCTCCCGACATGATGCTGAGCTGGAGACTA 250
181 nserGlyAsnSerPheLeuProAspValTyrMetLeuGlnLeuGln 198
251 GTTGGGGCCCAACATCCACGCAAGTCCAGCAGTTCAGGATGGCTTT 300
198 eTTPGGLYGLnGlnSerThrGlnValGlnHisSerThrAspGlyVal 214
301 AATAACCATTTGGTGGTGGTATATCCACCGGTCATCTCGTACGCT 350
215 AsnAsnProIleArgLeuAlaIleProGlyAsnPheValThrLeu 231
351 CAATGTTGGCAGCTGATCCGACCTTACGATCATGTTGTATGTA 400
231 AsnValAlaArgAspValIleAlaSerLeuAlaIleMetLeuPhe 248
401 GGCACCGCACATCTTCGCC 420
248 YgluArgProSerSerSer 254

seq_name: SwissProt_40:ABRA_ABRPR
seq_documentation_block:
ID ABRA_ABRPR STANDARD: PRT: 528 AA.
AC P1140: P28589.
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Abrin-a precursor [Contains: Abrin-a A chain (rRNA N-glycosidase)]
DE (EC 3.2.2.22): Abrin-a B chain.
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93132798; PubMed=8421313.
RA Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.;
RT "Primary structure of three distinct isoabins determined by cDNA
RT sequencing. Conservation and significance."
RL J. Mol. Biol. 229:263-267(1993).
RN [2]
RP SEQUENCE OF 1-251.
RC TISSUE=Seed;
RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;
RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic
RT protein from the seeds of Abrus precatorius."
RL Agric. Biol. Chem. 52:1095-1097(1988).
RN [3]
RP SEQUENCE OF 1-251 FROM N.A.
RC TISSUE=Leaf;

```

```

RX MEDLINE=91201329; PubMed=2016300;
RA Evensen G., Mathiesen A., Sundan A.;
RT "Direct molecular cloning and expression of two distinct abrin
RT A-chains."
RL J. Biol. Chem. 266:6848-6852(1991).
RN [4]
RP SEQUENCE OF 262-528.
RX MEDLINE=92371656; PubMed=1505674;
RA Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;
RT "The complete primary structure of abrin-a B chain."
RL FEBS Lett. 309:115-118(1992).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).
RX MEDLINE=95333188; PubMed=7608980;
RA Tanirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
RT "Crystal structure of abrin-a at 2.14 A."
RL J. Mol. Biol. 250:354-367(1995).
CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC ABRIN-A IS MORE TOXIC THAN RICIN.
CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
CC PRECEDES ENDOCYTOSIS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
CC -1- INACTIVATING FAMILY, TYPE 2 RIP SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M98344; AAA32624.1; ALT_INIT.
DR EMBL: X54872; -; NOT_ANNOTATED_CDS.
DR PIR: S2429; T2LSA.
DR PIR: S24133; S24133.
DR PDB: IABR; 07-FEB-95.
DR InterPro: IPR001574; RIP.
DR InterPro: IPR000772; Ricin_B_lectin.
DR Pfam: PF00652; Ricin_B_lectin; 2.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS50231; RICIN_B-LECTIN; 2.
DR PROSITE: PS00275; SHIGA-RICIN; 1.
KW Hydroxylase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;
KW Lectin; 3d-structure.
FT CHAIN 1 251 ABRIN-A A CHAIN.
FT PEPTIDE 252 261 LINKER PEPTIDE.
FT CHAIN 262 528 ABRIN-A B CHAIN.
FT DOMAIN 273 400 RICIN B-TYPE LECTIN 1.
FT DOMAIN 403 527 RICIN B-TYPE LECTIN 2.
FT REPEAT 283 325 1-ALPHA.
FT REPEAT 326 366 1-BETA.
FT REPEAT 369 401 1-GAMMA.
FT REPEAT 414 449 2-ALPHA.
FT REPEAT 453 492 2-BETA.
FT REPEAT 495 528 2-GAMMA.
FT ACT_SITE 164 164 BY SIMILARITY.
FT DISULFID 247 269 INTERCHAIN (BY SIMILARITY).
FT DISULFID 286 305 BY SIMILARITY.
FT DISULFID 329 346 BY SIMILARITY.
FT DISULFID 417 430 BY SIMILARITY.
FT DISULFID 456 473 BY SIMILARITY.

```



FT MOD\_RES 1 1 PYROLIDONE CARBOXYLIC ACID.  
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .).  
FT CONFLICT 202 202 MISSING (IN REF. 2).  
FT CONFLICT 298 298 M -> Y (IN REF. 4).  
FT CONFLICT 427 427 M -> L (IN REF. 4).  
FT CONFLICT 467 467 T -> P (IN REF. 4).  
FT CONFLICT 483 483 V -> L (IN REF. 4).  
SQ SEQUENCE 528 AA; 59243 MM; AIF76BEC5B9A827 CRC64;

alignment\_scores:  
Quality: 283.00 Length: 185  
Ratio: 2.320 Gaps: 6  
Percent Similarity: 65.946 Percent Identity: 38.919

alignment\_block:  
US-09-627-165B-15 x ABRA\_ABRPR ..

Align seg 1/1 to: ABRA\_ABRPR from: 1 to: 528

1 TCACAGATCTGGAGCATAGCCGGTCAT...AGGACGATCCGCT 47  
|||||  
113 TYGTYASPLSLGLAARGTTPALHSLGINSERARGLNGINLEP 129  
48 GGGTATAGAGAACTCATTCAGTCGCGCTCGGCTGTTATCCAGCG 97  
|||||  
129 UGLYGLNGLALGLLGLTHHGLYGLSERPHEARGSERGLYGL 146  
98 GCGACGACCGGGCCCAAGCTCGTCCCTATATCCCTGATTCAGATG 147  
|||||  
146 SNAAPSANGUGLULYSALARGHLEULLEVALLELLELIMETVAL 162  
148 TCCGAGGCGCGGATTCATCCATCTTTGGAGGCGTCCGCAATAC 197  
|||||  
163 ALGLNLALALARGPHEARGTYRILESERASARGVALARGVALSER 179  
198 TAACAGCGGGAGCATTTCTCCGACATGTACATCCGACGAGTCGAG 247  
|||||  
179 EGLTHHGLYTHRALPHEGLNPROASPALALAMETLIESERLEUGLA 196  
248 CTAGTTGGGGGACACATCCACGCAAGTCCAGAGCTTCAGGATGGCT 297  
|||||  
196 SNASTTPAPASANLSEARGLYVALGLNGLINSERVALGINASPTHR 212  
298 TTATAATACCATTCGTTGGGTATATCCACGCGTAACTCTGAGAGTT 347  
|||||  
213 PHEPROASNGLN.....VALTHRL 219  
348 GAGCAATGTTGCGGAC.....G 364  
|||||  
219 UTHRSNLLHARGANGLUPROVALLEVALASPERLEUSERHISPROT 236  
365 TGATGCGCAGCTTACGATCATGTTGTTGATATAGGACCGGACCAT 414  
|||||  
236 HVALALVALLEUALLEULMETLEUPHEVALYS...ASNPPOPSAN 251  
415 TCCCTCGAGCGTGGCTATTTGGCGGTACAGACCCGCTTGGA 464  
|||||  
252 ALASANGINSER.....PROLEULEHARGSERILEVALGLULY 265  
465 TAGCGCGCGCGTTCGACGATGTACCTGCATGCTCCGACACCGCG 514  
|||||  
265 SSER.....LYSILECYSERSEARGLYGLINPROTHVALA 278  
515 GCATC 519  
|||||  
278 RGLLE 279

seq\_name: swissprot\_40:RIC1\_RICCO

seq\_documentation\_block:  
ID RIC1\_RICCO STANDARD; PRT; 576 AA.

AC P02879; P02880;  
DT 21-JUL-1986 (rel. 01, Created)  
DT 13-AUG-1987 (rel. 05, Last sequence update)  
DT 01-MAR-2002 (rel. 41, Last annotation update)  
DE R1cin precursor [contains: R1cin A chain (rRNA N-glycosidase)  
DE (EC 3.2.2.22); R1cin B chain).  
OS R1cinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosida I; Malpighiales; Euphorbiaceae; R1cinus.  
OX NCBI\_TaxID=3986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86067214; PubMed=2999712;  
RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,  
RA Weaver R.F.;  
RT "Genomic cloning and characterization of a ricin gene from R1cinus  
RT communis.";  
RL Nucleic Acids Res. 13:8019-8033(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92163016; PubMed=1371405;  
RA Tregear J.W., Roberts L.M.;  
RT "The lectin gene family of R1cinus communis: cloning of a functional  
RT ricin gene and three lectin pseudogenes.";  
RL Plant Mol. Biol. 18:515-525(1992).  
RN [3]  
RP SEQUENCE OF 12-576 FROM N.A.  
RX MEDLINE=85179479; PubMed=3838723;  
RA Lamb A., Roberts L.M., Lord J.M.;  
RT "Nucleotide sequence of cloned cDNA coding for preproricin.";  
RL Eur. J. Biochem. 148:265-270(1985).  
RN [4]  
RP SEQUENCE OF 36-302.  
RA Yoshitake S., Funatsu G., Funatsu M.;  
RT "Isolation and sequences of peptic peptides, and the complete  
RT sequence of the chain of ricin-D.";  
RL Agric. Biol. Chem. 42:1267-1274(1978).  
RN [5]  
RP SEQUENCE OF 315-576.  
RA Funatsu G., Kimura M., Funatsu M.;  
RT "Primary structure of Ala chain of ricin D.";  
RL Agric. Biol. Chem. 43:2221-2224(1979).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.  
RX MEDLINE=90344223; PubMed=1368517;  
RA Kimura Y., Kusoku H., Tada M., Takagi S., Funatsu G.;  
RT "Structural analyses of sugar chains from ricin A-chain variant.";  
RL Agric. Biol. Chem. 54:157-162(1990).  
RN [7]  
RP REVIEW.  
RX MEDLINE=21480122; PubMed=11595634;  
RA Olshes S., Kozlov J.V.;  
RT "Ricin.";  
RL Toxicon 39:1723-1728(2001).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
RX MEDLINE=87165983; PubMed=3558397;  
RA Monfort W., Villafraña J.E., Monzingo A.F., Ernst S.R., Katzin B.,  
RA Rutenber E., Xiong N.H., Hamlin R., Robertus J.D.;  
RT "The three-dimensional structure of ricin at 2.8 A.";  
RL J. Biol. Chem. 262:5396-5403(1987).  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
RX MEDLINE=91352004; PubMed=1881881;  
RA Katzin B.J., Collins E.J., Robertus J.D.;  
RT "Structure of ricin A-chain at 2.5 A.";  
RL Proteins 10:251-259(1991).  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.  
RX MEDLINE=91352005; PubMed=1881882;  
RA Rutenber E., Robertus J.D.;  
RT "Structure of ricin B-chain at 2.5-A resolution.";

DR	PDB:	1IFS;	14-JAN-98.
DR	PDB:	1IEU;	14-JAN-98.
DR	PDB:	1IOB;	31-OCT-93.
DR	PDB:	1OBS;	16-JUN-97.
DR	PDB:	1OBT;	16-JUN-97.
DR	PDB:	1BR5;	02-SEP-98.
DR	PDB:	1BR6;	02-SEP-98.
DR	GlycoSiteDB:	P02879..	-
DR	InterPro:	IPR001574,	RIP.
DR	InterPro:	IPR000772;	RicIn_B_lectin.
DR	pfam:	PF00652;	RicIn_B_lectin.
DR	pfam:	PF00161;	RIP; 1.
DR	PRINTS:	PR00396;	SHIGARICIN.
DR	SMART:	SM00458;	RICIN; 2.
DR	PROSITE:	PS50231;	RICIN_B_LECTIN; 2.
KW	PROSITE:	PS00275;	SHIGA-RICIN; 1.
KW	Hydrolase:	Protein synthesis inhibitor;	Toxin; Repeat; Glycoprotein;
FT	Lectin:	Signal;	3D-structure.
FT	SIGNAL	1	35
FT	CHAIN	36	302
FT	PEPTIDE	303	314
FT	CHAIN	315	576
FT	DOMAIN	321	448
FT	DOMAIN	451	575
FT	REPEAT	331	373
FT	REPEAT	374	414
FT	REPEAT	417	449
FT	REPEAT	462	497
FT	REPEAT	501	540
FT	REPEAT	543	570
FT	ACT_SITE	212	212
FT	DISULFID	294	318
FT	DISULFID	334	353
FT	DISULFID	377	394
FT	DISULFID	465	478
FT	DISULFID	504	521
FT	CARBOHYD	45	45
FT	CARBOHYD	271	271
FT	N-LINKED (GLCNAC. . .)	/FTid-CAR_000080.	.
FT	N-LINKED (GLCNAC. . .); IN MINOR A-CHAIN VARIANT.		
FT	N-LINKED (GLCNAC. . .)	/FTid-CAR_000081.	
FT	N-LINKED (GLCNAC. . .)		
FT	N-LINKED (GLCNAC. . .)		
FT	E -> D (IN REF. 3).		
FT	A -> R (IN REF. 3).		
FT	TURN	43	47
FT	TURN	49	50
FT	HELIX	53	67
FT	STRAND	73	74
FT	TURN	75	76
FT	STRAND	77	79
FT	TURN	88	90
alignment_scores:			
	Quality:	259.50	Length: 178
	Ratio:	2.257	Gaps: 5
Percent Similarity:	64.607	Percent Identity:	41.011
alignment_block:			
US-09-627-1658-15 x RICL_RICCO ..			
Align seg 1/1 to: RICL_RICCO from: 1 to: 576			
1 TAAACAATCTGAGCGATTAGCCGGTCA...AGGACACAGACCNCNT 47		:::      :::    :::	
158 TYASPARGLEUGlueGlueAlaGlaYAsnLeuArgGluAsnIleGluLe 174		::::::::::::::::::::	
48 GGGATATAGGAACACTCATTCACGCCGTCTGCAGCTTGTTAT..... 90			
174 ugluyangngylProleugnglugluAalaleterAtaleuatyrytyrIsert 191			

```

between the Swiss Institute of Bioinformatics and the EMBL outstation at its European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or\_send\_an\_email\_to\_licenselsb-ch).
-----
CC      EMBL; X55667; CAI39202.1; -
DR      PIR; S16022; S16022.
DR      HSSP; P11140; IABR.
DR      InterPro; IPRO01574; RIP.
DR      InterPro; IPRO00772; Ricin_B_lectin.
DR      Pfam; PF00652; Ricin_B_lectin; 2.
DR      Pfam; PF00161; RIP; 1.
DR      PRINTS; PRO0396; SHICARICIN.
DR      SMART; SM00458; RICIN_2_LECTIN; 2.
DR      ProSITE; PS50231; SHICA_RICIN; 1.
KW      Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;
KW      Lectin; Signal.
FT      SIGNAL          1       34
FT      CHAIN           35       285
FT      PEPTIDE         286       295
FT      CHAIN           296       562
FT      DOMAIN          307       434
FT      REPEAT          317       359
FT      REPEAT          360       400
FT      REPEAT          403       435
FT      REPEAT          448       483
FT      REPEAT          487       526
FT      REPEAT          529       562
FT      ACT_SITE        198       198
FT      DISULFID        281       303
FT      DISULFID        320       339
FT      DISULFID        363       380
FT      DISULFID        451       464
FT      DISULFID        480       507
FT      MOD_RES         35
SQ      CARBOHYD        234       234
FT      CARBOHYD        335       395
FT      CARBOHYD        435       435
SQ      SEQUENCE        562 AA; 62817 MW; 1FD0ABC7D7BA6278 CRC64;

alignment_scores:
Quality: 256.00      Length: 175
Ratio: 2.151        Gaps: 5
Percent Similarity: 68.000    Percent Identity: 40.000

alignment_block:
US-09-627-165B-15 x ABRC_ABRR ..

Align seg 1/1 to: ABRC_ABRR from: 1 to: 562

1 TACACAGATCTGAGCGATAGCCGGTCAT..AGGACCAGATCCCTCT 47
||| |||||||||||:::||| ::| ||:::||| ||
147 TygIAsprLeugIArgTrpAlaHisGlnThrArgIgluIleSerIe 163
48 GGGTATAGAGCACTCATTCATCCGTCGCCGCTTGTTCACAGGCG 97
||||:::||| ||| :::::::||| ||||| ::||
163 uclIleuGlnAlaLeuThrIshAlaIleSerPheLeuArgseGlyAlas 180
98 GCAGCACC GCCGCCAGACTGTTCCCTTAATTATCTCATTCAGATGATC 147
::: :: ::|||::|::|::|::|::|::|::|::|
180 eTasnspcIugluysAlaargThreuleValleIlelleGlnetala 196
148 TCAGAGCGCGGAGATTCATCCCTTTTGAGGGCTGCCCATATCAT 197
||||| |||||||::|::|::| ||| ||::| |
197 SerGIAlaIalargTytraGtyrIleSerAshmrgValdyIalSerIl 213
198 TAACACCGGCGAGTCATTTCTCCGACATGTACATCTCGACCTGAGA 247

```



CC CC INACTIVATING FAMILY. TYPE 2. RIP SUBFAMILY. -----  
 CC CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS. -----  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC CC the European Bioinformatics Institute. There are no restrictions on its  
 CC CC use by non-profit institutions as long as its content is in no way  
 CC CC modified and this statement is not removed. Usage by and for commercial  
 CC CC entities requires a license agreement (see [http://www.isb-sib.ch/announce/or\\_send\\_an\\_email\\_to\\_license@isb-sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch)).  
 CC CC -----  
 DR EMBL: M98345; AAA32625.1; -  
 DR HSSP: P11401; IABR.  
 DR InterPro: IPR001574; RIP.  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 2.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW Hydroxylase; protein synthesis inhibitor; Toxin; Repeat; glycoprotein;  
 KW Lectin.  
 FT CHAIN 1 250 ABRIN-B A CHAIN.  
 FT PEPTIDE 251 260 LINKER PEPTIDE.  
 FT CHAIN 261 527 ABRIN-B B CHAIN.  
 FT DOMAIN 272 399 RICIN B-TYPE LECTIN 1.  
 FT DOMAIN 402 526 RICIN B-TYPE LECTIN 2.  
 FT REPEAT 282 324 1-ALPHA.  
 FT REPEAT 325 365 1-BETA.  
 FT REPEAT 368 400 1-GAMMA.  
 FT REPEAT 413 448 2-ALPHA.  
 FT REPEAT 452 491 2-BETA.  
 FT REPEAT 494 527 2-GAMMA.  
 FT ACT\_SITE 163 163 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 246 268 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 285 304 BY SIMILARITY.  
 FT DISULFID 328 345 BY SIMILARITY.  
 FT DISULFID 416 429 BY SIMILARITY.  
 FT DISULFID 435 472 BY SIMILARITY.  
 FT MOD\_RES 1 1 PYROLIDONE CARBOXYLIC ACID (BY SIMILARITY).  
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 282 282 N -> D (IN REF. 2).  
 FT CONFLICT 291 291 D -> N (IN REF. 2).  
 FT CONFLICT 350 351 AE -> PQ (IN REF. 2).  
 FT CONFLICT 351 351 S -> N (IN REF. 2).  
 FT CONFLICT 378 378 Y -> M (IN REF. 2).  
 FT CONFLICT 426 426 Y -> D (IN REF. 2).  
 FT CONFLICT 428 428 R -> S (IN REF. 2).  
 FT CONFLICT 431 431 R -> K (IN REF. 2).  
 FT CONFLICT 434 434 N -> S (IN REF. 2).  
 FT CONFLICT 484 484 H -> Y (IN REF. 2).  
 FT CONFLICT 491 491 H -> Q (IN REF. 2).  
 FT CONFLICT 493 493 E -> G (IN REF. 2).  
 FT CONFLICT 502 502 H -> M (IN REF. 2).  
 FT CONFLICT 509 509 H -> W (IN REF. 2).  
 FT CONFLICT 513 513 H -> T (IN REF. 2).  
 FT CONFLICT 516 516 H -> T (IN REF. 2).  
 SQ SEQUENCE 527 AA; 59114 MW; 3253AE490CE9494A CRC64;

alignment\_scores: 246.50 Length: 175  
 Quality: 2.071 Gaps: 6  
 Ratio: 2.071 Percent Identity: 38.857

Percent Similarity: 68.000

alignment\_block:  
 US-09-627-165B-15 x ABRB\_ABRPR ..  
 Align seg 1/1 to: ABRB\_ABRPR from: 1 to: 527

1 TACACAGATCTGCAGCGATACGCCGTCAT...AAGACACGATCCCTCT 47

```

RT      "Isolation and partial characterization of nigrin b, a non-toxic
RT      novel type 2 ribosome-inactivating protein from the bark of Sambucus
RT      nigra L.,"
RL      Plant Mol. Biol. 22:1181-1186(1993).
CC      -1- FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN
CC      PROTEIN SYNTHESIS BUT DOES NOT AFFECT PLANT NOR BACTERIAL PROTEIN
CC      SYNTHESIS. THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC      SUBUNITs THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC      B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE
CC      BINDING OF NIGRIN B TO THE CELL MEMBRANE THAT PRECEDES
CC      ENDOCYTOSIS.
CC      -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC      specific adenosine on the 28S rRNA.
CC      -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC      -1- SIMILARITY: IN THE N-TERMINAL SUBDOMAIN, BELONGS TO THE RIBOSOME-
CC      INACTIVATING FAMILY, TYPE 2 RIP SUBFAMILY.
CC      -1- SIMILARITY: CONTAINS 2 RCIN B-TYPE LECTIN DOMAINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC      or send an email to license@isb-slb.ch).
CC      -----
DR      EMBL; U041299; AAB39475.1; -.
DR      PIR; S37382; S37382.
DR      PIR; S37383; S37383.
DR      HSSP; P02879; 2AAT.
DR      Mendel; 15504; Sammi.R1p.15504.
DR      InterPro: IPR000772; Ricin_B_lectin.
DR      InterPro: IPR001574; RIP.
DR      Pfam; PF00652; Ricin_B_lectin; 2.
DR      Pfam; PF00161; RIP; 1.
DR      PRINTS; PRO0396; SHIGARICIN.
DR      SMART; SM00458; RICIN; 2.
DR      PROSITE; PS00275; SHIGA_RICIN; 1.
DR      PROSITE; PS02311; RICIN_B_LECTIN; 2.
KW      Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;
KW      lectin; signal.
FT      SIGNAL          1      25
FT      CHAIN           26      297
FT      CHAIN           298      563
FT      DOMAIN          305      431
FT      DOMAIN          434      559
FT      REPEAT          316      356
FT      REPEAT          357      397
FT      REPEAT          400      432
FT      REPEAT          445      482
FT      REPEAT          486      524
FT      REPEAT          527      554
FT      REPEAT          563      597
FT      ACT_SITE        188      188
FT      DISULFID        274      302
FT      DISULFID        319      338
FT      DISULFID        360      377
FT      DISULFID        448      463
FT      DISULFID        489      506
FT      CARBOHYD        221      221
FT      CARBOHYD        368      368
FT      CARBOHYD        376      376
FT      CARBOHYD        483      483
FT      CARBOHYD        537      537
FT      CONFLICT        39      39
SQ      SEQUENCE        563 AA; 62300 MW; F250CBE24621BF14 CMC64;

```



227 ThrAgtGlyGlyGluPheAlaArgProValGluLeuArgThrValSerAs 243  
327 CACCGTACTCTGCTGACGTCGAGCAATGTCGGAGCGATGCCAGCT 376  
243 nThrProThrPheValThrAsnValAsn...SerProValValysGlyI 259  
377 TAGCATATATGTTGTTTCTAT.....GTAGGACGACCATCTT 415  
259 leuAlaLeuLeuLeuLeuYr.PheArgValAlaValGlyThrAspAsnVal 274  
seq\_name: SwissProt\_40:RIP1\_TRIKI  
seq\_documentation\_block:  
ID RIP1\_TRIKI STANDARD: PRT: 289 AA.  
AC P09989;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Ribosome-inactivating protein alpha-trichosanthin precursor  
DE (rRNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS).  
OS Trichosanthes Kirilowii (Mongolian snake-gourd).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
OX NCBI\_TaxID=3677;  
[1]  
RN RP SEQUENCE FROM N.A.  
RC STRAIN-MAXIMOWICZ;  
RX MEDLINE=91153657; PubMed=1999291;  
RA Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;  
RT "Cloning of trichosanthin cDNA and its expression in Escherichia  
RT coli.";  
RL Gene 97:267-272(1991).  
[2]  
RN RP SEQUENCE FROM N.A.  
RC STRAIN-MAXIMOWICZ; TISSUE=leaf;  
RX MEDLINE=90256790; PubMed=2341400;  
RA Chow T., Feldman R.A., Lovett M., Pataak M.;  
RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a  
RT type I ribosome-inactivating protein.";  
RL J. Biol. Chem. 265:8670-8674(1990).  
[3]  
RN RP SEQUENCE OF 24-270.  
RC STRAIN-MAXIMOWICZ; TISSUE=tubercous root;  
RX MEDLINE=90256789; PubMed=2341399;  
RA Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R.,  
RT Wu P., Hwang K., Pataak M.;  
RT "Primary amino acid sequence of alpha-trichosanthin and molecular  
RT models for abrin A-chain and alpha-trichosanthin.";  
RL J. Biol. Chem. 265:8665-8669(1990).  
[4]  
RN RP SEQUENCE OF 24-270.  
RC TISSUE=tubercous root;  
RA Wang Y., Qian R.O., Gu Z.W., Jin S.W., Zhang L.O., Xia Z.X.,  
RT Tian G.Y., Ni C.Z.;  
RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and  
RT application.";  
RL Pure Appl. Chem. 58:789-798(1986).  
[5]  
RN RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).  
RX MEDLINE=94344957; PubMed=8066085;  
RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;  
RT "Structure of trichosanthin at 1.88-A resolution.";  
RL Proteins 19:4-13(1994).  
RN RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
RX MEDLINE=95344383; PubMed=7619070;  
RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;  
RT "Studies on crystal structures, active centre geometry and  
RT deprotecting mechanism of two ribosome-inactivating proteins.";  
RL Biochem. J 309:285-298(1995).  
CC -1- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS  
CC CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT

CC CC INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
CC specific adenosine on the 28S rRNA.  
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1  
CC RIP SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb.ch/announce/>  
CC or send an email to [license@isb.ch](mailto:license@isb.ch)).  
CC -----  
DR EMBL: M34858; AAA34207.1; -;  
DR EMBL: J05434; AAA34206.1; -;  
DR PIR: J70003; RLTRP.  
DR PIR: J70566; J70566.  
DR PIR: A36273; A36273.  
DR PIR: A36274; A36274.  
DR PDB: IMR; 07-FEB-95.  
DR PDB: IMR; 07-FEB-95.  
DR PDB: IMR; 10-JUL-95.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARCIN.  
DR PROSITE: PS00273; SHIGARCIN; 1.  
KW Antitumor; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;  
KW 3D-structure.  
FT SIGNAL 1 23  
FT CHAIN 24 270  
FT PROPEP 271 289  
FT ACT SITE 183 183  
FT CONFLICT 57 60  
FT CONFLICT 82 84  
FT CONFLICT 87 87  
FT CONFLICT 92 92  
FT CONFLICT 143 144  
FT CONFLICT 196 196  
FT CONFLICT 214 216  
FT CONFLICT 231 231  
FT CONFLICT 234 234  
FT CONFLICT 246 266  
FT CONFLICT 247 247  
SQ SEQUENCE 289 AA; 31676 MW; 5CE09B630575B89 CRC64;  
  
alignment\_scores:  
Quality: 163.50 Length: 170  
Ratio: 1.587 Gaps: 5  
Percent Similarity: 60.588 Percent Identity: 30.000  
  
alignment\_block:  
US-09-627-165B-15 x RIP1\_TRIKI ..  
Align seg 1/1 to: RIP1\_TRIKI from: 1 to: 289  
  
1 TACACAGATCTGAGCGATACGCGGTCAAT...AGGACGATCCCTCT 47  
111 ||||| |||||  
134 TyrgluatrgluenglnhrAlaAlaGlyLysIleatrglulsmIleProLe 150  
48 GGTATAGAGACTCATTCATCCGCTCTGGCGCTTCGTTATCCAGGCG 97  
111 ||||| |||||  
150 uGlyLeuProAlaLeuAspserAlaIlethrThrLeuPheTyrTyrAsn 167  
98 GCACACCGCGCGCCAGCTGCTTCCTTATATCCCTCATTCAGATGATC 147  
111 ||||| |||||  
167 laaAsnserAlaAlaIleAsnAla.....LeuMetValLeuIleGlnserThr 181  
148 TCCGAGCGCGGAGATTCATTCATCCATCTTTGGAGGCGCTCGCATATAC 197  
111 ||||| |||||  
182 SerGlnAlaAlaIleTyrIlyspheIleGlnGlnIleGlyIlyAsArgVa 198





RC	TISSUE=Seed;
RX	MEDLINE-91242096; PubMed-1366643;
RA	Kung S.S., Kimura M., Funatsu G.;
RT	"The complete amino acid sequence of antiviral protein from the seeds
RT	of pokeweed ( <i>Phytolacca americana</i> ).";
RL	Agric. Biol. Chem. 54:3301-3318(1990).
CC	-1- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN
CC	SYNTHESIS IN VITRO.
CC	-1- CATALYTIC ACTIVITY: Endopolylysis of the N-glycosidic bond at on
CC	specific adenosine on the 28S rRNA.
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY, TYPE 1
CC	RIP SUBFAMILY.
DR	PIR; J06401; J06401.
DR	HSSP; Q03464; IAPA.
DR	InterPro; IPRO01574; RIP.
DR	Pfam; PF00161; RIP; 1.
DR	PRINTS; PR00396; SHIGARICIN.
DR	PROSITE; PS00275; SHIGA_RICIN; 1.
KM	Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin.
FT	ACT. SITE 175 175
FT	DISULFD 34 258
FT	DISULFD 84 105
SC	SEQUENCE 261 AA; 29200 MW; D88B99962FE8399D CRC64;

Align seg 1/1 to: RIPS\_PHYAM from: 1 to: 261

[illegible]

seq\_name: SwlssProt\_40:RIPA\_PHYAM

seq\_documentation\_block:

ID	STANDARD:	PRF:	294 AA.
AC	RIBA_PHYM		
AC	003464;		
DT	01-OCT-1993 (Rel. 27, Created)		
DT	01-OCT-1993 (Rel. 27, Last sequence update)		
DT	01-MAR-2002 (Rel. 41, Last annotation update)		
DE	Antiviral protein alpha precursor (PAP-alpha)		(Ribosome-Inactivating
DE	protein) (TRNA N-glycosidase) (EC 3.2.22).		
OS	Phytolacca americana (common pokeweed)		(Virginian pokeweed).
OC	Eukaryota: Viridiplantae; Streptophyta;		Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicts;		
OC	Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.		
OC	NCBI_TaxId=3527;		

RN SEQUENCE FROM N.A.  
 RN TISSUE=Seed, Leaf, and Root;  
 RX MEDLINE=93099240; PubMed=128138;  
 RA Katooka J., Habuka N., Masuta C., Miyano M., Koizai A.;  
 RT Isolation and analysis of a genomic clone encoding a pokeweed  
 RN plant protein. J. Biol. Chem. 267:879-886(1992).  
 RN Plant Mol. Biol. 20:879-886(1992).

ANAL CRISTALLOGRAPHY (2.3 ANGSTROMS)  
MEDLINE=95010127; Pubmed=7925458;

RA Ayo H., Kataoka J., Tsuge H., Hakuba N., Inagaki E., Noma M.,  
Miyano M.;

RA Miyano M.

"X-ray structure of a pokeweed antiviral protein, coded by a new genomic clone, at 0.23-nm resolution. A model structure provides a suitable electrostatic field for substrate binding." ;  
Eur. J. Biochem. 225:369-374(1994).  
+1-FTJGCEQO;TWYZZL

-1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 specific adenosine on the 28S rRNA.  
 -1- SUBUNIT: MONOMER

-1- SUBUNIT: MONOMER.

1- SURCELLULAR LOCATION: Cell wall.  
1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY, TYPE 1.  
RIP SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/-----send-an-email-to-licensing@isb-sib.ch>).

EMBL; D10600; BAA01451.1; -

PIR; S28421; S28421.

PDB; 1APA; 31-JAN-94.

InterPro; IPR001574; RIP

PF00161; RIP; 1.

PROSITE; PS00275; SHIGA\_

Antiviral; Protein synth

Cell wall; 3D-structure.

SIGNAL	1	24
T		

CHAIN	25	285
-------	----	-----

T	PROPER	286	294
---	--------	-----	-----

ACT_SITE	199	199
----------	-----	-----

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```

FT DISULFID 58 282
FT DISULFID 108 130
FT STRAND 28 31
FT HELIX 32 34
FT HELIX 37 51
FT STRAND 57 58
FT STRAND 59 60
FT STRAND 61 63
FT TURN 67 68
FT STRAND 73 79
FT TURN 81 82
FT STRAND 85 91
FT TURN 92 94
FT STRAND 97 104
FT TURN 105 106
FT STRAND 107 113
FT TURN 114 115
FT HELIX 119 129
FT STRAND 136 138
FT STRAND 141 141
FT HELIX 147 154
FT TURN 155 155
FT HELIX 158 160
FT STRAND 163 163
FT HELIX 165 175
FT TURN 176 177
FT HELIX 183 195
FT TURN 196 197
FT HELIX 198 202
FT HELIX 204 212
FT TURN 213 215
FT STRAND 218 218
FT HELIX 222 240
FT STRAND 242 242
FT TURN 243 244
FT STRAND 245 253
FT TURN 255 256
FT STRAND 259 264
FT HELIX 265 271
FT STRAND 275 275
SQ SEQUENCE 294 AA: 33069 MW: F2EC27724FA85596 CRC64:

```

```

alignment_scores:
  Quality: 156.00      Length: 134
  Ratio: 1.660         Gaps: 3
  Percent Similarity: 70.149   Percent Identity: 30.597

```

alignment\_block:  
US-09-627-165b-15 x RIPA\_PHYAM ..

Align seg 1/1 to: RIPA\_PHYAM from: 1 to: 294

```

1 TACACAGATCTGGAGCAGTACGCCGCTCAT..AGGACACAGATCCCTT 47
   ||| |||||: :|||: :|||: :|||: |||
147 TProAlaLeuGluLysValGlyArgProArgSerGlnValGlnLe 163
48 GCGTATGAGGACTCATTCATCGCTCGCGCTTCGTTATCCAGGCG 97
   |||||: :||| :|||: :|||: :|||: :|||
163 uGlyLeuGlnLeuAsnSerGlyIleGlyLysIleTyrGlyValAsp 180
98 GCACACACC...CGGCCCAAGCTCGTTCCTTAATCCTCATTCAGATG 144
   :||| :|||: :|||: :|||: :|||: :|||
180 erPheThrGluLysThrGluAlaGluPheLeuValAlaIleGlnMet 196
145 ATCTCCAGGCGCGAGATTCATCCATCCATCTTTGGAGGCGCCAAAT 194
   :|||: :|||: :|||: :|||: :|||: :|||
197 ValSerGluAlaAlaArgPheLysTyrIleGluAsnGlnValLys... 211
195 CATTAACGCGGAGATCTTCTCCGACATGTACATGCTCGAGCTGG 244
   ||| :|||: :|||: :|||: :|||: :|||
212 .ThrAsnPheAsnArgAlaPheTyrProAsnAlaLysValLeuAsnLeu 288

```

```

245 AGACTAGTGGGGCCAAATCCAGCCAGTCCAGCATGCTACGATGCGC 294
   || |||||: :|||: :|||: :|||: :|||
228 LuGluSerTrpGlyIleSerThrAlaIleHisAsnAlaLysAsnGly 244
295 GTTTTATATACCATTCGTTGGGTATATCCACCGGTAACCTGTCAC 344
   :|||: :|||: :|||: :|||: :|||: :|||
245 AlaLeuThrSerProLeuGluLeuLysAsnAlaLysGlySerLysTrp 261
345 GTTGAGCATGTTCCGACGATCGCCAGCTTACGATCATGTTGTTG 394
   :|||: :|||: :|||: :|||: :|||: :|||
261 eValLeuArgValAspAspIleGluProAspValGlyLeuLeuLysTyr 278
395 TA 396
   ||
278 al 278

seq_name: SwissProt_40:RIP1_MOMCH

seq_documentation_block:
ID RIP1_MOMCH STANDARD: PRT: 286 AA.
AC P16094; P24697;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAR-1992 (Rel. 21, last sequence update)
DT 01-MAR-2002 (Rel. 41, last annotation update)
DE Ribosome-inactivating protein momordin I precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22) (Alpha-momorcharin) (Alpha-MMC).
OS Momordica charantia (Bitter melon) (Balsam pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_Taxid=3673;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=91159486; PubMed=2001404;
RA Ho W.K.K., Liu S.C., Shaw P.C., Yeung H.W., Ng T.B., Chan W.Y.;
RT "Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating
RT protein."
RL Biochim. Biophys. Acta 1088:311-314(1991).
RN [2]
RP SEQUENCE OF 24-38.
RC TISSUE=Seed;
RX MEDLINE=89326691; PubMed=2753596;
RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,
RA Lappi D.;
RT "N-terminal sequence of some ribosome-inactivating proteins."
RL Int. J. Pept. Protein Res. 33:263-267(1989).
RN [3]
RP SEQUENCE OF 24-70.
RC TISSUE=Seed;
RX MEDLINE=89005108; PubMed=3262509;
RA Casellas P., Dussossoy D., Falasca A.I., Barbieri L., Stirpe F.,
RA Guillemot J.C., Ferrara P., Bolognesi A., Cenini P., Stirpe F.;
RT "Trichokirin, a ribosome-inactivating protein from the seeds of
RT Trichosanthes kirilowii Maximowicz. Purification, partial
RT characterization and use for preparation of immunotoxins."
RL Eur. J. Biochem. 176:581-588(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=94356447; PubMed=8075985;
RA Ren J., Wang Y., Dong Y., Stuart D.I.;
RT "The N-glycosidase mechanism of ribosome-inactivating proteins
RT implied by crystal structures of alpha-momorcharin."
RL Structure 2:7-16(1994).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).
RX MEDLINE=94192822; PubMed=8143869;
RA Husain J., Tickle I.J., Wood S.P.;
RT "Crystal structure of momordin, a type I ribosome inactivating
RT protein from the seeds of Momordica charantia."
RL FEBS Lett. 342:154-158(1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).

```

```

RX MEDLINE-95344383; PubMed=7619070;
RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT "Studies on crystal structures, active-centre geometry and
RL dephosphorylating mechanism of two ribosome-inactivating proteins.";
CC Biochem. J. 309:285-298(1995).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY, TYPE 1
CC R1P SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL; X57682; CAA04869.1; -.
DR PIR; S14273; RLPUG.
DR PIR; S16490; S16490.
DR PDB; 1AHN; 22-JUN-94.
DR PDB; 1AHB; 22-JUN-94.
DR PDB; 1AHC; 22-JUN-94.
DR PDB; 1AMW; 31-MAY-94.
DR PDB; 1MRG; 07-FEB-95.
DR PDB; 1MRH; 07-FEB-95.
DR PDB; 1MRJ; 07-FEB-95.
DR GlycosultedB; P16094; -.
DR InterPro; IP001574; R1P.
DR Pfam; PF00161; R1P; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KM Protein synthesis inhibitor; Hydrolase; Toxin; Signal; Glycoprotein;
KW 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 269 RIBOSOME-INACTIVATING PROTEIN MOMORDIN I.
FT PROPEP 270 286 MISSING IN MATURE PROTEIN.
FT ACT_SITE 183 183
FT CARBOHYD 250 250 N-LINKED (GLICNAC. . .).
FT FTID 286 286 /FTID-CAR_000082.
SQ SEQUENCE 286 AA; 31532 MW; E1B013ABEC216CF CRC64;

alignment_scores:
    Quality: 154.50      Length: 135
    Ratio: 1.736        Gaps: 4
    Percent Similarity: 65.926      Percent Identity: 32.593

alignment_block:
US-09-627-165b-15 x R1P1_MOMCH ..

Align seg 1/1 to: R1P1_MOMCH from: 1 to: 286

1 TACAGAGATCTGGAGGATACGCGGTCAT...AGGACCATGCTCT 47
   ||| |||::: ||||| |||::: |||::: |||::: |||:::
134 TygIaIrgLeuGInIleAlaIaGlyLysProArgGlnLysIleProI1 150
48 GGGTATAGAGAACTCATTCATCGCTCGGCGCTGCTTATCCAGCG 97
   |||::: ||| |||::: |||::: |||::: |||::: |||:::
150 eGIyleuProAlaIeuAspSerAlaIleSerThrIleuLeuHis.....T 165
98 GGAGCACCGCGGCCAGAGCTGCTCCCTTATATATCCATTCAGATG 147
   ::::: ||| ||| |||::: |||::: |||::: |||::: |||:::
165 yRAspSerThrAlaIaIaGlyAlaIaLeuValIleuIleGlnThrThr 181
148 TCCGAGCGCGGAGATCATTCATCCATTTGGAGGGCTCCCATATAC 197
   ::::: ||| ||| |||::: |||::: |||::: |||::: |||:::
182 AlaGluAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 198
198 TAAACAGCGGGAGATCTTCTCCGACATGTACATGCTCAGATGAGA 247
   ||| ||| |||::: |||::: |||::: |||::: |||::: |||:::
198 aTyraIrgAspGlu.....ValProSerLeuAlaThrIleSerLeuGln 213

```

```

seq_name: SwissProt_40:R1P1_BRD1
seq_documentation_block:
ID R1P1_BRD1 STANDARD; PRT; 290 AA.
AC P33185; 09S819;
DT 01-OCT-1993 (Rel. 27, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein bryodin 1 precursor (rRNA N-glycosidase)
DE (EC 3.2.2.22) (BDI).
OS Bryonia dioica (Red bryony).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosid I; Cucurbitales; Cucurbitaceae; Bryonia.
OX NCBI_TaxID=3652;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RC TISSUE=leaf;
RX MEDLINE-97228081; PubMed-9115985;
RA Gawlak S.L., Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,
RA Siegall C.B.;
RT "Molecular, biological, and preliminary structural analysis of
RT recombinant bryodin 1, a ribosome-inactivating protein from the plant
RT Bryonia dioica."
RL Biochemistry 36:3095-3103(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Siegall C.B.;
RT "Cloning and expression of a gene encoding bryodin 1 from Bryonia
RT dioica."
RL Patent number US5541110, 30-JUL-1996.
RN [3]
RP SEQUENCE OF 24-66.
RC TISSUE=Seed;
RX MEDLINE-89326691; PubMed-2753596;
RA Montecucchi P.C., Lazzerini A.M., Barbieri L., Stipe F., Soria M.,
RA Lappi D.;
RT "N-terminal sequence of some ribosome-inactivating proteins."
RL Int. J. Pept. Protein Res. 33:263-267(1989).
RN [4]
RP SEQUENCE OF 24-43.
RC TISSUE=Root;
RX MEDLINE-95151812; PubMed-7849072;
RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,
RA Marguardt H.;
RT "Characterization of ribosome-inactivating proteins isolated from
RT Bryonia dioica and their utility as carcinoma-reactive
RT immunconjugates."
RL Bioconj. Chem. 5:423-429(1994).
CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS
CC PROTEIN SYNTHESIS IN ANIMAL CELLS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- PTM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO
CC PRODUCE A SHORTER PROTEIN.
CC -1- BIOTECHNOLOGY: Especially useful as immunotoxin for

```

pharmacological applications as it has low toxicity in rats and mice but is potent once inside target cells.  
-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1 RIP SUBFAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL: 124020; -; NOT\_ANNOTATED\_CDS.  
DR PIR: S16491; S16491.  
DR PDB: 1BRV; 04-MAR-98.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
KW Protein synthesis inhibitor; Hydrolase; Toxin; 3D-structure;  
KW Multigene family; Glycoprotein; Signal.  
FT SIGNAL 1 23  
FT CHAIN 24 270 RIBOSOME-INACTIVATING PROTEIN BRYODIN I.  
FT PROPEP 271 290 MISSING IN MATURE PROTEIN.  
FT ACT\_SITE 183 183 BY SIMILARITY.  
FT ACT\_SITE 212 212  
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT MUTAGEN 212 212 E->K: REDUCES ACTIVITY 10-FOLD.  
FT CONFLICT 61 65 RSSIS -> LRHXI (IN REF. 3).  
SQ SEQUENCE 290 AA; 31788 MW; E966CD9C031A42DB CRC64;

alignment\_scores:  
Quality: 145.50 Length: 135  
Ratio: 1.565 Gaps: 5  
Percent Similarity: 68.889 Percent Identity: 31.111

alignment\_block:

US-09-627-165b-15 x RIP1\_BRYDI ..

Align seg 1/1 to: RIP1\_BRYDI from: 1 to: 290

```

1 TACACAGATCTGGAGCGATACGCCGATCAT..AGGACGACGATCCCTCT 47
   ||| ||||| ||||| ||| ||||| |||||
134 TyrcLunrgrleuGlnThrAlaAlaGlyLysIleargLusnleProle 130
   ||| ||||| ||||| ||||| ||||| |||||
48 GGTATAGAGAACTCAATCCGCTCGCGCGCTTCGTTATCCAGCG 97
   ||||| ||||| ||||| ||||| ||||| |||||
150 uGlyLeuProAlaLeuAspSerAlaIleThrLeuTyTyrThra 167
   ||||| ||||| ||||| ||||| ||||| |||||
98 GCACACCGCGCGCCAGCTGCTCCCTTAAATCCATTCGATGATC 147
   ||||| ||||| ||||| ||||| ||||| |||||
167 lAsSerSerAlaAlaSerAla.....LeuLeuValLeuIleGlnSerThr 181
   ||||| ||||| ||||| ||||| ||||| |||||
148 TCCGAGCCGCGAGATTCATCCCATCTTTGGAGCGCTGCCAATACAT 197
   ||||| ||||| ||||| ||||| ||||| |||||
182 AlagLueRAlaargTyrlsPheIleGlnGlnGlnIleGlyLysArgVa 198
   ||||| ||||| ||||| ||||| ||||| |||||
198 TAAACAGCGGGGATCATTTCTCCGACATGATACATCTCGAGCTGAGA 247
   ||||| ||||| ||||| ||||| ||||| |||||
198 lAsp.....LysThrPheLeuProSerLeuAlaThrIleSerLeuAlu 213
   ||||| ||||| ||||| ||||| ||||| |||||
248 CTAGTTGGGCGCACATCCAGCATGTCAG.....CAGTCAGGAT 291
   ||||| ||||| ||||| ||||| ||||| |||||
213 snAsnTrpSerAlaLeuSerLysIleGlnIleAlaSerThrAsnAsn 229
   ||||| ||||| ||||| ||||| ||||| |||||
292 GCGCTTTTAAATACCATTCGTTGGGTATATACACGCGTAATCTCGT 341
   ||||| ||||| ||||| ||||| ||||| |||||
230 GlylnPheGlnSerProValIleLeuIleAspLysnAsnGlnArgVa 246
   ||||| ||||| ||||| ||||| ||||| |||||
342 GACGTTGACGCAAT.....GTTCCGCGACGTGATCGCACGCTTACGATCA 385

```

```

||||| ||| ||| ||||| |||||
246 lSerIleThrAsnAlaSerAlaArgValValThrSerAsnIleAlaLeuL 263
366 TGTGTG 390
|||||
263 euleu 264

```

seq\_name: SwissProt\_40:RIP1\_PHYAM

seq\_documentation\_block:

```

ID RIP1_PHYAM STANDARD; PRT; 313 AA.
AC P10297;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Antiviral protein I precursor (PAP-I) (PAP-C) (Ribosome-inactivating
DE protein) (tRNA N-glycosidase) (EC 3.2.2.22).
GN PAP1.
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.
OX NCBI_TaxID=3527;
[1]
RN RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-62.
RP TISSUE=Leaf;
RC MEDLINE=92003676; PubMed=1912488;
RA Lin Q., Chen Z.C., Antoniw J.F., White R.F.;
RT "Isolation and characterization of a cDNA clone encoding the
RT anti-viral protein from Phytolacca americana.";
RL Plant Mol. Biol. 17:609-614(1991).
[2]
RN RP SEQUENCE OF 23-65.
RP MEDLINE=69193489; PubMed=2930487;
RA Barbieri L., Bolognesi A., Centini P., Falasca A.I., Minghetti A.,
RA Garofano L., Guicciardi A., Lappi D., Miller S.P.;
RT "Ribosome-inactivating proteins from plant cells in culture.";
RL Biochem. J. 257:801-807(1989).
[3]
RN RP SEQUENCE OF 23-54.
RP TISSUE=Leaf;
RC MEDLINE=83290867; PubMed=6885760;
RA Houston L.L., Ramakrishnan S., Hermodson M.A.;
RT "Seasonal variations in different forms of pokeweed antiviral protein,
RT a potent inactivator of ribosomes.";
RL J. Biol. Chem. 258:9601-9604(1983).
[4]
RN RP SEQUENCE OF 23-54.
RP TISSUE=Leaf;
RC MEDLINE=85023392; PubMed=6091760;
RA Bjorn M.J., Larrick J., Platak M., Wilson K.J.;
RT "Characterization of translocated inhibitors from Phytolacca
RT americana. amino-terminal sequence determination and antibody-
RT inhibitor conjugates.";
RL Biochim. Biophys. Acta 790:154-163(1984).
[5]
RN RP SEQUENCE OF 23-54.
RP TISSUE=Root;
RC MEDLINE=91064383; PubMed=2248976;
RA Bolognesi A., Barbieri L., Abbondanza A., Falasca A.I., Carniceilli D.,
RA Battelli M.G., Stirpe F.;
RT "Purification and properties of new ribosome-inactivating proteins
RT with RNA N-glycosidase activity.";
RL Biochim. Biophys. Acta 1087:293-302(1990).
[6]
RN RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RP MEDLINE=94016586; PubMed=8411176;
RA Monzinger A.F., Collins E.J., Ernst S.R., Irvin J.D., Robertus J.D.;
RT "The 2.5 A structure of pokeweed antiviral protein.";
RL J. Mol. Biol. 233:705-715(1993).
[7]
RN RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 23-284.
RP MEDLINE=99421320; PubMed=10493577;

```

RA Kurinov I.V., Myers D.E., Irvin J.D., Uckun F.M.;  
 FT "X-ray crystallographic analysis of the structural basis for the  
 FT interactions of pokeweed antiviral protein with its active site  
 FT inhibitor and ribosomal RNA substrate analogs.";  
 RL Protein Sci. 8:1765-1772(1999)  
 CC -1- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN  
 CC SYNTHESIS IN VITRO.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- TISSUE SPECIFICITY: SPRING LEAVES AND ROOTS.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY, TYPE 1  
 CC RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X53583; CAA39054.1; -;  
 DR PIR; S02792; S02792.  
 DR PIR; S13469; S13469.  
 DR PDB; 1PAF; 31-JAN-94.  
 DR PDB; 1PAG; 31-JAN-94.  
 DR PDB; 1QCI; 15-SEP-99.  
 DR PDB; 1QCG; 15-SEP-99.  
 DR PDB; 1QCF; 15-SEP-99.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 DR Activity: Protein synthesis inhibitor; Hydrolase; Toxin; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 22  
 FT CHAIN 23 285  
 FT PROPEP 286 313  
 FT ACT\_SITE 198 198  
 FT DISULFID 56 281  
 FT DISULFID 107 128  
 FT STRAND 25 29  
 FT HELIX 35 49  
 FT STRAND 55 56  
 FT STRAND 57 58  
 FT STRAND 59 61  
 FT TURN 65 66  
 FT STRAND 71 78  
 FT HELIX 79 81  
 FT STRAND 82 89  
 FT TURN 90 93  
 FT STRAND 94 102  
 FT TURN 103 105  
 FT STRAND 106 112  
 FT TURN 113 114  
 FT TURN 117 117  
 FT HELIX 118 127  
 FT TURN 131 132  
 FT STRAND 134 136  
 FT STRAND 139 139  
 FT HELIX 145 152  
 FT TURN 153 153  
 FT TURN 157 158  
 FT STRAND 162 162  
 FT HELIX 164 174  
 FT TURN 175 176  
 FT HELIX 182 195  
 FT TURN 196 196  
 FT HELIX 197 201  
 FT TURN 203 211  
 FT STRAND 212 214  
 FT STRAND 217 217  
 FT HELIX 221 239

ANTIVIRAL PROTEIN I.  
 BY SIMILARITY.

SO	SEQUENCE	313 AA;	35219 MW;	2C57B2861EBA57F5 CRC64;
FT	STRAND	241	241	
FT	TURN	242	243	
FT	STRAND	244	252	
FT	TURN	254	255	
FT	STRAND	258	263	
FT	HELI	264	270	
FT	STRAND	274	274	

alignment\_scores:  
 Quality: 144.50 Length: 135  
 Ratio: 1.606 Gaps: 3  
 Percent Similarity: 66.667 Percent Identity: 30.370

alignment block:  
 US-09-627-165b-15 x RIP1\_PHYM ..

Align seq 1/1 to: RIP1\_PHYM from: 1 to: 313

```

1 TACACAGATCTGGAGCGATACGGCGGT.....CATAGGACCAAGATCCC 44
||| ||||| ||||| ||||| |||||
145 TYTProthirleuGlusSerLysAlaGlyValLysSerArgSerGlnValGI 161
45 TCTGGATATAGAGGAGACTCATTCATCCGCTCGCGCTTCGTTATCCAG 94
||| ||||| ||| :||| :||| :||| :||| :||| :||| :||| :|||
161 nleuGlyIleGlnIleleuaspSerasnIleGlyLysIleSerGlyValM 178
95 GCGGCGACACC...CGGCGCCAGCTCGTCCCTTATATCCATCCAG 141
:|: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
178 etSerPheThrIleGlyThrGluAlaGluPheLeuLeuValAlaIleGln 194
142 ATGATCTCCGAGCGCGAGATTCATTCATCCATCTTTGGAGGCGCTGGCA 191
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
195 MetValSerGlnAlaAlaArgPheLysTYRlleGluasnGlnValLys.. 210
211 ....ThrAsnPheasnArgAlaPheasnProasnProLysValLeuAsnL 226
242 TGGAGACTAGTTGGGCGCAACATCCAGCAAGTCAGACGTACGGAT 291
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
226 euGlnGluThrIleGlyLysIleSerThrAlaIleHisAspAlaLysasn 242
292 GCGGTTTTTATTAACCAATTCGTTGGGATATCCAGCGGTAACTTCGT 341
||| ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
243 GlyValLeuProLysProLeuGlnLeuValAspAlaSerGlyAlaLysTr 259
342 GACGTTGAGCAATGTTGCGAGCGTGAATCCAGACTTAGCATGATTCGT 391
:|: :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
259 pIleValleuArgValaspGlnIleLysProaspValAlaLeuLeuAsnT 276
392 TNGTA 396
276 yVal 277

```

Wed Jul 31 08:46:46 2002

us-09-627-165b-15.rspt

Page 1

OM of: US-09-627-165b-15 to: SPREMBL\_19.\* out-format : pfs

Date: Jul 30, 2002 4:05 PM

About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL-frame+np2.model -DEV-xlp  
-O/cpnp2.1/USPNC.spool/US09627165/rnat.30072002.151140.3298/app_query.fasta_1.1012  
-DS-SPREMBL_19 -QRM-fastan -SUPPLY-rspt -GAPOP-12.000  
-GAPOP-4.000 -MINMATCH-0.100 -LOOPEX-0.000 -LOOPEXT-0.000  
-GAPOP-4.500 -GAPOP-0.050 -XGAPOP-10.000 -XGAPOP-0.500  
-GAPOP-6.000 -GAPOP-7.000 -YGAPOP-10.000 -YGAPOP-0.500  
-DELOP-6.000 -DELEXT-7.000 -START-1 -MATRIX-blosum62  
-TRANS-human40.cdi -LIST-45 -DOCCALIGN-200 -THR_SCORE-pct  
-THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL -OUTFMT-pfs  
-NORM-ext -HEAPSIZE-500 -MINLEN-2000000000  
-USER-US09627165 -CGN1_1.310 -NCPU-6 -ICPU-3 -LONGLOG  
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -NO_XLPHY -WAIT -THREADS-1
```

Search information block:

Query: US-09-627-165b-15

Query length: 522

Database: SPREMBL\_19.\*

Database sequences: 562222

Database length: 17294929

Search time (sec): 124.830000

score list:

Sequence	Strd	Orig	2Score	EScore	len	Document
SP-plant:094BW3	+	319.50	612.30	1.6e-26	580	094bw3 cinnaomum campora (can
SP-plant:094BW4	+	317.50	608.30	2.7e-26	580	094bw4 cinnaomum campora (can
SP-plant:094BW5	+	309.50	592.89	2.0e-25	549	094bw5 cinnaomum campora (can
SP-plant:094BW6	+	309.50	592.28	2.1e-25	581	094bw6 cinnaomum campora (can
SP-plant:094BW7	+	259.50	493.20	7.5e-20	541	094bw7 cinnaomum campora (can
SP-plant:094BW8	+	256.00	486.31	1.8e-19	528	094bw8 cinnaomum campora (can
SP-plant:094BW9	+	249.00	480.24	8.4e-19	525	094bw9 cinnaomum campora (can
SP-plant:094BW10	+	243.00	468.28	3.9e-18	521	094bw10 cinnaomum campora (can
SP-plant:094BW11	+	242.00	466.28	5.0e-18	521	094bw11 cinnaomum campora (can
SP-plant:094BW12	+	237.00	447.48	2.5e-17	570	094bw12 cinnaomum campora (can
SP-plant:094BW13	+	236.00	445.86	2.3e-17	570	094bw13 cinnaomum campora (can
SP-plant:094BW14	+	232.00	435.48	9.1e-17	604	094bw14 cinnaomum campora (can
SP-plant:094BW15	+	231.00	435.48	1.2e-16	570	094bw15 cinnaomum campora (can
SP-plant:094BW16	+	225.50	424.58	4.7e-16	565	094bw16 cinnaomum campora (can
SP-plant:094BW17	+	222.00	417.93	1.1e-15	547	094bw17 cinnaomum campora (can
SP-plant:094BW18	+	220.00	402.24	1.4e-15	552	094bw18 cinnaomum campora (can
SP-plant:094BW19	+	216.50	406.24	1.7e-15	569	094bw19 cinnaomum campora (can
SP-plant:094BW20	+	201.00	375.60	2.5e-13	564	094bw20 cinnaomum campora (can
SP-plant:094BW21	+	193.00	359.56	2.0e-12	563	094bw21 cinnaomum campora (can
SP-plant:094BW22	+	192.50	357.88	2.3e-12	563	094bw22 cinnaomum campora (can
SP-plant:094BW23	+	189.50	357.62	2.5e-12	563	094bw23 cinnaomum campora (can
SP-plant:094BW24	+	189.50	357.62	2.5e-12	563	094bw24 cinnaomum campora (can
SP-plant:094BW25	+	183.00	339.62	3.8e-12	563	094bw25 cinnaomum campora (can
SP-plant:094BW26	+	183.00	339.62	2.5e-11	563	094bw26 cinnaomum campora (can
SP-plant:094BW27	+	175.00	334.37	2.6e-11	563	094bw27 cinnaomum campora (can
SP-plant:094BW28	+	163.50	331.44	1.4e-10	298	094bw28 cinnaomum campora (can
SP-plant:094BW29	+	160.50	307.77	2.9e-09	289	094bw29 cinnaomum campora (can
SP-plant:094BW30	+	153.50	286.91	4.0e-08	313	094bw30 cinnaomum campora (can
SP-plant:094BW31	+	152.00	284.29	5.7e-08	302	094bw31 cinnaomum campora (can
SP-plant:094BW32	+	151.50	284.44	6.2e-08	270	094bw32 cinnaomum campora (can
SP-plant:094BW33	+	149.00	281.45	1.0e-07	247	094bw33 cinnaomum campora (can
SP-plant:094BW34	+	142.00	265.14	7.2e-07	279	094bw34 cinnaomum campora (can
SP-plant:094BW35	+	141.00	263.14	9.3e-07	279	094bw35 cinnaomum campora (can
SP-plant:094BW36	+	141.00	262.88	9.4e-07	279	094bw36 cinnaomum campora (can
SP-plant:094BW37	+	137.50	254.88	2.4e-06	314	094bw37 cinnaomum campora (can
SP-plant:094BW38	+	136.50	255.88	2.8e-06	314	094bw38 cinnaomum campora (can
SP-plant:094BW39	+	136.00	252.88	3.4e-06	286	094bw39 cinnaomum campora (can
SP-plant:094BW40	+	135.00	251.74	4.3e-06	264	094bw40 cinnaomum campora (can
SP-plant:094BW41	+	128.00	233.84	2.7e-05	315	094bw41 cinnaomum campora (can

SP-plant:000980	+	123.50	228.18	8.3e-05	278	000980 luffa cylindrica (smo
SP-plant:098805	+	122.00	225.49	0.0001	270	098805 amarantus viridis. r
SP-plant:096322	+	119.00	218.95	0.0003	284	096322 amarantus tricolor (
SP-plant:099418	+	116.50	214.41	0.0005	272	099418 beta vulgaris (sugar
SP-plant:093y65	+	111.50	203.58	0.0018	294	093y65 dianthus chinensis. r

seq\_name: SP-plant:094BW3

seq\_documentation\_block:

ID	NAME	PRELIMINARY	PRT	580 AA.
AC	094BW3			
DT	01-DEC-2001 (Trembl)	19	Created	
DT	01-DEC-2001 (Trembl)	19	Last sequence update	
DT	01-DEC-2001 (Trembl)	19	Last annotation update	
DE	TYPE 2 RIBOSOME-INACTIVATING PROTEIN CINNAMOMIN III			
DE	PRECURSOR.			
OS	Cinnamomum camphora (Camphor tree).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.			
OX	NCBI_TaxID=13429;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Yang Q., Gong Z.Z., Liu W.Y.			
RT	"Molecular cloning of three type 2 RIP (ribosome-inactivating protein)			
RT	genes encoding cinnamomin proteins and study of their expression			
RT	patterns."			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY039803; AKR82460.1;			
KW	Signal.			
FT	SIGNAL	1	32	POTENTIAL.
FT	CHAIN	33	580	TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT	SEQUENCE	580 AA; 64421 MW; 940D10F01EFB558 CRC64;		

alignment\_scores:

Quality: 319.50 Length: 184

Percent Similarity: 72.826 Percent Identity: 46.196

alignment\_block:

US-09-627-165b-15 x 094BW3

Align seg 1/1 to: 094BW3 from: 1 to: 580

1	TACACGATCGGAGGATCGGCGGT...	CATAGGACGATCGCTT	47
147	TYTTHASPLDGLUATGVALAASJYGLUATGATGUGUILLLEULE		163
48	GGGTATGAGGACTCATTCATATCCTCGGCGCTTCATTACGAGC		96
163	UGLYETASPROLEUATLUSALATLESERIALLEUTRPILESERSL		180
97	..GGCAGCAGCCGCGGCAAGCTCGTCCCTATATATCTCATTCAGATG		144
180	EUASGLGALNARGALALEUALARGSERLEULLLEIVALLGLEMET		196
145	ATCTCGAGGCGGAGATCATCATCTTGGAGAGGCTGCCATA		194
197	VALAIAAGLUALAVALATGHEATGHEILLEGUITYFATGVALATGGLUSE		213
195	CATTATACGCGGAGATCATTCCTTCCGACATGATTCATCGATCG		244
213	RTLEHTRARGALALUETHATGHEATGHEATGHEATGHEATGHEATG		230
245	ACACTATGTTGGCGCAATTCATTCACGACGATTCACGATTCAGAT		291
230	LUASLYETSPSERALALEUSERASALALVALICGLISERASGLNGLY		246
292	GGCGTTTATATACCATTCGTTGGATATATACGCGGATTCATTCGT		341
247	GLYVALPHESERPROVALGULUATGATGATGATGATGATGATGATG		263





Percent Similarity: 71.739 Percent Identity: 45.652

alignment\_block:  
US-09-627-165B-15 x Q9FV22 ..

Align seg 1/1 to: Q9FV22 from: 1 to: 549

```

1  TACACGATCTGAGCGATACGCCGT...CATAGGACCAATCCCTCT 47
   |||||||
115  TYRTHASPLEUGLGLYVALALAGLYGUAARGGLUGLILEU 131
   |||||||
48  GGGTATAGAGAACTCATTCATCCGCTCGCGCTTATCCAGGC 96
   |||||||
131  UGLYMETASPROLEUGLUSNAALALESERALALEUTRPILESERANL 148
   |||||||
97  ..GGCAGCACCCGGCCCAAGCTGCTTATATATCCATTCAGATG 144
   |||||||
148  EUNANGINGLARGALALEUALARGSERLILEVALILEGLINMET 164
   |||||||
145  ATCTCGAGAGCGCGAGATTCACATCTTTGGAGGCGTCGCCAATA 194
   |||||||
165  VALALAGLUALVALARGPHEARGPHEILEGLUTRARGVALARGLYSE 181
   |||||||
195  CATTAACAGCGGGAGTCATTCTCCCGCATGTCATGTCGAGCTGG 244
   |||||||
181  RILESERARGALAGLUMETPHEARGPROASPROALAMELLEUSERLEUG 198
   |||||||
245  AGACTAGTTGGGGCCCAACATCCAGCAAGTCACAGCTCTACGGAT... 291
   |||||||
198  LUSANLYSTRPSERLALALEUSERASNALVALGLINGINSEASNLGLY 214
   |||||||
292  GCGGTTTTTAATACCATTCCTGGTGGTATATCCAGCGTAATCTGCT 341
   |||||||
215  GLYVALPHESESERPROVALGLULEUARGSERLILESERANLYSPROVA 231
   |||||||
342  GACGTGGACCAATGTTCCGAC...GTGATCGCGAGCTTAGGAGCAATG 388
   |||||||
231  ITRYVALIGLYSERVALSERASPARGVALILESERGLYLEUALILEMETL 248
   |||||||
389  TGTTTGTATGTAG...GACCCACCATCTTCTCCGACGTCGCCAT 432
   |||||||
248  EUPHELLECYASRGSETRHASPARGLASERSERASPLINPHEILEASP 264
   |||||||
433  TGGCCCTGGTGCATACGACCCGCTCTGAAAT...AG 467
   |||||||
265  HISMETLEUMETLLEARGPROLILEUVALASPVALALAGLUALALATH 281
   |||||||
468  CGGCGCGCGTCGAGATGTACCTGACGCTCCGAAACCGACGCGGCGCA 517
   |||||||
281  RASPALASAPASNASPAPRINCYSALIASAPROGLUPROTHRVALARGI 298
   |||
518  TC 519
   ||
298  LE 298

```

seq\_name: sp\_plant:Q94BWS

seq\_documentation\_block:

AC Q94BWS; PRELIMINARY; PRT; 581 AA.  
 DT 01-DEC-2001 (TREMBLREL. 19, Created)  
 DT 01-DEC-2001 (TREMBLREL. 19, Last sequence update)  
 DE TYPE 2 RIBOSOME-INACTIVATING PROTEIN CINNAMOMIN I PRECURSOR.  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 RX NCBI\_TaxId=13429;  
 RN 111  
 RP SEQUENCE FROM N. A.  
 RA Yang Q., Gong Z.Z., Liu W.X.;  
 RT Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 genes encoding cinnamomin proteins and study of their expression

RT patterns:"  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY039801; AAK83458.1; ..  
 KW SIGNAL.  
 FT SIGNAL 1 32 POTENTIAL.  
 FT CHAIN 33 581 TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIN I.  
 SQ SEQUENCE 581 AA: 64215 MW: 6585F8BFBA3D196 CRC64:

alignment\_scores:

Quality: 309.50 Length: 184  
 Ratio: 2.345 Gaps: 6  
 Percent similarity: 71.739 Percent identity: 45.652

alignment\_block:  
US-09-627-165B-15 x Q94BWS ..

Align seg 1/1 to: Q94BWS from: 1 to: 581

```

1  TACACGATCTGAGCGATACGCCGT...CATAGGACCAATCCCTCT 47
   |||||||
147  TYRTHASPLEUGLGLYVALALAGLYGUAARGGLUGLILEU 163
   |||||||
48  GGGTATAGAGAACTCATTCATCCGCTCGCGCTTATCCAGGC 96
   |||||||
163  UGLYMETASPROLEUGLUSNAALALESERALALEUTRPILESERANL 180
   |||||||
97  ..GGCAGCACCCGGCCCAAGCTGCTTATATATCCATTCAGATG 144
   |||||||
180  EUNANGINGLARGALALEUALARGSERLILEVALILEGLINMET 196
   |||||||
145  ATCTCGAGAGCGCGAGATTCACATCTTTGGAGGCGTCGCCAATA 194
   |||||||
197  VALALAGLUALVALARGPHEARGPHEILEGLUTRARGVALARGLYSE 213
   |||||||
195  CATTAACAGCGGGAGTCATTCTCCCGCATGTCATGTCGAGCTGG 244
   |||||||
213  RILESERARGALAGLUMETPHEARGPROASPROALAMELLEUSERLEUG 230
   |||||||
245  AGACTAGTTGGGGCCCAACATCCAGCAAGTCACAGCTCTACGGAT... 291
   |||||||
230  LUSANLYSTRPSERLALALEUSERASNALVALGLINGINSEASNLGLY 246
   |||||||
292  GCGGTTTTTAATACCATTCCTGGTGGTATATCCAGCGTAATCTGCT 341
   |||||||
247  GLYVALPHESESERPROVALGLULEUARGSERLILESERANLYSPROVA 263
   |||||||
342  GACGTGGACCAATGTTCCGAC...GTGATCGCGAGCTTAGGAGCAATG 388
   |||||||
263  ITRYVALIGLYSERVALSERASPARGVALILESERGLYLEUALILEMETL 280
   |||||||
389  TGTTTGTATGTAG...GACCCACCATCTTCTCCGACGTCGCCAT 432
   |||||||
280  EUPHELLECYASRGSETRHASPARGLASERSERASPLINPHEILEASP 296
   |||||||
433  TGGCCCTGGTGCATACGACCCGCTCTGAAAT...AG 467
   |||||||
297  HISMETLEUMETLLEARGPROLILEUVALASPVALALAGLUALALATH 313
   |||||||
468  CGGCGCGCGTCGAGATGTACCTGACGCTCCGAAACCGACGCGGCGCA 517
   |||||||
313  RASPALASAPASNASPAPRINCYSALIASAPROGLUPROTHRVALARGI 330
   |||
518  TC 519
   ||
330  LE 330

```

seq\_name: sp\_plant:Q41174

seq\_documentation\_block:

AC Q41174; PRELIMINARY; PRT; 541 AA.  
 RT Q41174;



```

98 GCAGACACCGGGCCAAAGGCTTCCTCCCTTAAATCCATTAAGTATGATC 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
146 cctsmatpoclugluyalsalatrghtrheullevalllelbelmeca1a 162
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
148 TCCGAGCGCGGAGATTCATCCCATCTTTGGAGGCGCGCAATACAT 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 Seegluablaalatrghtrhcylllebrasmrvalglvalserll 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
198 TAACAGCGCGGAGTCATTTCTTCCGACATGTACATGCTCGAGCTGAGA 247
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 earghrghllyrhlabheglrProApPProAlaMetleuSerleuGua 196
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
248 CTAGTTGGGGCCCAACATCCAGCAGCAAGTCAGCAGAGTCTACGGATGGCGT 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
196 snbsnrtPAspAsnleuSerGlyglvalGlnGlnInservalGlnAspAla 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
298 TTT...AATAACCCATTTGGTGGGGTATATCCACCGGTAACCTTGTCAC 344
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
213 PheProAnaenValIleleuSerSerIleAsnArglnProValaIva 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
345 GTTAGACAAATGTTGGGACGTATGATCCGAGGTATAGCAGATGTTGTTG 394
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
229 lAspSerleuSerhIAspThrValAlaValleuAlaIalaMetleuPheV 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
395 TAGGTAGGAGCGACACATCTCTCTCGACAGTGGCGGATATGGCGCGTGC 444
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
246 aIcys...AsnProAnaenAlaIasnInsr.....ProleuLeu 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
445 ATACGACCCGCTGTGGAAANATAGCGGCGCGCTGACAGATTTACCTGCAC 494
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
259 lEaHysSerIleValGlnGlnSer.....LysIcysSer 271
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
495 TGGTTCGACACCCACCGTGGCCATC 519
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
271 rArgTyrGlnuProThrValArgIle 279
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: sp_plant:O38761

seq_documentation_block
ID O38761 PRELIMINARY; PRT; 252 AA.
AC O38761; O96234;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE rRNA - GLYCOSIDASE (EC 3.2.2.22) (FRAGMENT).
OS RIP.
NC Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Kosidae;
OC eucotsids I; Fabales; Fabaceae; Papilionoideae; Adreae; Abrus;
OC NCBI_TaxId:3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAVES;
RA Evensen G., Mathiesen A., Sundan A.;
RL Submitted (OCT-1990) to the EMBL/Genbank/DBJ databases.
[2]
RP SEQUENCE OF 2-252 FROM N.A.
RC TISSUE=SEED;
RA MEDLINE-94139756; PubMed-8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur J Biochem. 219:83-87(1994).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 26S RNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
DR EMBL: X54873; CA38655.1; -.
DR EMBL: X76720; CA54138.1; -.
DR RSSP: P1140; IABR.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP; 1.
DR

```

[illegible]

RT "Cloning and expression of three abrin A-chains and their mutants  
RT derived by site-specific mutagenesis in *Escherichia coli*.";  
RL Eur. J. Biochem. 219:83-87(1994).  
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.  
DR EMBL: X76721; CAA54139.1; -.  
DR HSSP: P11140; 1ABR.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
KW Hydrolase; Toxin.  
FT NON\_TER 1 1  
FT NON\_TER 251 251  
SQ SEQUENCE 251 AA; 28040 MW; D57ECB182E0EC9 CMC64;

## alignment\_scores:

Quality:	243.00	Length:	145
Ratio:	2.305	Gaps:	3
Percent Similarity:	66.897	Percent Identity:	40.000

## alignment\_block:

US-09-627-165B-15 x Q96236 ..

Align seg 1/1 to: Q96236 from: 1 to: 251

```
1 TACACAGATCTGAGCGATACCGCGGTGAT...AGGACGACAGATCCCTCT 47
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 TYRGLYSPLRLEUGLARGTRPALHISGLINSEARGINGLILEPROLE 129
48 GGGTATAGAGGAGCTCATTCATCCGTCGCGGCGGCTTATCCGATACCGAGCG 97
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
129 UGLYLEUGLINALLEUTHRHISGLYILESERPHEPHARGSERGLYGLYA 146
98 GAGACGCCGGGCGCAAGCTGTCCTTATATCCATTCAGATGAGATGAC 147
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
146 SNAEPASNGIUGLWYSALAARGTRHLEULEILEGIMETVAL 162
148 TCCGAGCGCGGAGATTCAATCCCATCTTTGGAGGCGCTGCCAATACAT 197
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 ALAALAIALAIALARGPHEARGYRILESERASNAARGVALARGVALSERIL 179
198 TACAGCGGGGATCATTTCTCCGACATGTACATCTGACGCTGAGAGA 247
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 EGLINHRGLYTHRALAPHEGLNPROASPALAIALAMELILESERLEUGLUA 196
248 CTACTTGGGGCCAAACATCCACGCAAGCTCCAGAGCTACGATGGCGCTT 297
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 SNAENITRPASPASNLEUSERARGLYVALGINGLINUSERVALGINASPTHR 212
298 TTATATTAACCCATTTGCGTTGGGTATATCCACCGGTAACTTCGTGACGTT 347
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
213 PHEPROASNGLN.....ValThrLe 219
348 GAGCAATGTCGCGAC.....G 364
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 UTHRASNILEARGASNGLUPROVALILEVALASPSERLEUSERHISPROT 236
365 TGATCGCGCAGCTTAGCGATCATGTTGTTGTATGCT 399
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 hVALAIAValLEuAlAlaLeuMetLeuPheValGys 247
```

seq\_name: sp\_plant:Q96237

seq\_documentation\_block:  
ID Q96237 PRELIMINARY; PRT; 251 AA.  
AC Q96237;  
DT 01-FEB-1997 (TREMBLREL. 02, Created)  
DT 01-FEB-1997 (TREMBLREL. 02, Last sequence update)  
DT 01-DEC-2001 (TREMBLREL. 19, Last annotation update)  
DE RRNA - GLYCOSIDASE (EC 3.2.2.22) (FRAGMENT).  
OS Abrus precatorius (Indian licorice) (Crab's eye).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosid 1; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
OK NCBI\_TaxID=3816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISUBSE-SEED;  
RA MEDLINE=94139756; PubMed=8307038;  
RX Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;  
RT "Cloning and expression of three abrin A-chains and their mutants  
RT derived by site-specific mutagenesis in *Escherichia coli*.";  
RL Eur. J. Biochem. 219:83-87(1994).  
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.  
DR EMBL: X76722; CAA54140.1; -.  
DR HSSP: P11140; 1ABR.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
KW Hydrolase; Toxin.  
FT NON\_TER 1 1  
FT NON\_TER 251 251  
SQ SEQUENCE 251 AA; 28055 MW; 6F64755C3DEAFB79 CRC64;

## alignment\_scores:

Quality:	242.00	Length:	145
Ratio:	2.495	Gaps:	3
Percent Similarity:	66.897	Percent Identity:	40.000

## alignment\_block:

US-09-627-165B-15 x Q96237 ..

Align seg 1/1 to: Q96237 from: 1 to: 251

```
1 TACACAGATCTGAGCGATACCGCGGTGAT...AGGACGACAGATCCCTCT 47
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 TYRGLYSPLRLEUGLARGTRPALHISGLINSEARGINGLILEPROLE 129
48 GGGTATAGAGGAGCTCATTCATCCGTCGCGGCGGCTTATCCGATACCGAGCG 97
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
129 UGLYLEUGLINALLEUTHRHISGLYILESERPHEPHARGSERGLYGLYA 146
98 GACACGCCGGGCGCAAGCTGTCCTTATATCCATTCAGATGAGATGAC 147
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
146 SNAEPASNGIUGLWYSALAARGTRHLEULEILEGIMETVAL 162
148 TCCGAGCGCGGAGATTCAATCCCATCTTTGGAGGCGCTGCCAATACAT 197
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 ALAALAIALAIALARGPHEARGYRILESERASNAARGVALARGVALSERIL 179
198 TACAGCGGGGAGATCATTTCTCCGACATGTACATCTGACGCTGAGAGA 247
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 EGLINHRGLYTHRALAPHEGLNPROASPALAIALAMELILESERLEUGLUA 196
248 CTACTTGGGGCCAAACATCCACGCAAGCTCCAGAGCTACGATGGCGCTT 297
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 SNAENITRPASPASNLEUSERARGLYVALGINGLINUSERVALGINASPTHR 212
298 TTATATTAACCCATTTGCGTTGGGTATATCCACCGGTAACTTCGTGACGTT 347
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
213 PHEPROASNGLN.....ValThrLe 219
348 GAGCAATGTCGCGAC.....G 364
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 UTHRASNILEARGASNGLUPROVALILEVALASPSERLEUSERHISPROT 236
365 TGATCGCGCAGCTTAGCGATCATGTTGTTGTATGCT 399
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 hVALAIAValLEuAlAlaLeuMetLeuPheValGys 247
```

seq\_name: sp\_plant:Q41358

```

seq_documentation_block:
ID 041358      PRELIMINARY;      PRT;      570 AA.
AC 041358:
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, last annotation update)
DE RNA - GLYCOSIDASE PRECURSOR (EC 3.2.2.22).
OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Adoxaceae; Sambucus.
OX NCBI_TaxID=4202;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BARK;
RX MEDLINE=96202926; PubMed=8611319;
RA Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;
RT "The Neucalalpha-2,6)-Gal/galNAc-binding lectin from elderberry
RT (Sambucus nigra) bark, a type-2 ribosome-inactivating protein with an
RT unusual specificity and structure.";
RL Eur. J. Biochem. 235:128-137(1996).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
DR EMBL; U27122; AAC49158.1; -.
DR HSSP; P02879; ZAAI.
DR InterPro; IPR000772; R1c1n_B_lectin.
DR Pfam; PF00652; R1c1n_B_lectin; 2.
DR Pfam; PF00161; R1P; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS0231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; signal; toxin.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 308 LECTIN_A CHAIN.
FT CHAIN 309 570 LECTIN_B CHAIN.
FT CHAIN 570 AA; 63101 MW; A059E2A3E86B868 CRC64;
SQ

```

Alignment\_scores:                      Length:                      186  
Quality:                      237.00                      Gaps:                      7  
Percent Similarity:                      66.667                      Percent Identity:                      34.946

Alignment\_block:  
US-09-627-165B-15 x 041358                      ..

Align seg 1/1 to: 041358 from: 1 to: 570

```

1 TACACGATCTGTGAGCGATACGGCGTCAT..AGGACACGATCCCTCT 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
145 TTTTSSerLeuGlnArgGlnValGlyPheGlyArgValTyrIleProLe 161
48 GGGTATAGAGAACTCATTCATCCGCTCTGGCGCTTCT.....T 88
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
161 uGlyProLysSerLeuAspGlnAlaIleSerSerLeuArgThrTyrTrpL 178
89 ATTCAGCGCGACAGCCCGCGGCCCAACCTGCTCCCTTAATTCCTCAT 138
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
178 euThrAlaGlyAspThrLysProLeuAlaArgGlyLeuLeuValValIle 194
139 CAGATGATCTCCGAGCGCGCGAGATTCATCCATCTTTTGGAGGGCTCT 188
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
195 GlnMetValSerGlnValAlaIleArgPheArgTyrIleGluLeuArgIle 211
189 CCATTAATTAACGCGCGGAGTCAATTCCTCCGACGATGACGCTCG 238
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
211 gThSerIleThrAspAlaSerGlnPheThrProAspLeuLeuMetLeu 228
239 AGCTGGAAGCTAGTTGGGCGCAACATCCAGCAAGTCCAGACAGTCT... 285

```

```

seq_name: sp_plant:Q96235
seq_documentation_block:
ID 096235      PRELIMINARY;      PRT;      251 AA.
AC 096235:
DT 01-FEB-1997 (TRENBLREL. 02, Created)
DT 01-FEB-1997 (TRENBLREL. 02, last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, last annotation update)
DE RNA - GLYCOSIDASE (EC 3.2.2.22) (FRAGMENT).
OS Arabus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
DR EMBL; X76644; CA54092.1; -.
DR HSSP; P11140; IABR.
DR InterPro; IPR001574; R1P.
DR Pfam; PF00161; R1P; 1.
DR PRINTS; PR00396; SHIGARICIN.
KW Hydrolase; Toxin.
FT NON_TER 1 251
FT NON_TER 251 1
FT NON_TER 251 1
SQ

```

Alignment\_scores:                      Length:                      145  
Quality:                      236.00                      Gaps:                      3  
Percent Similarity:                      66.207                      Percent Identity:                      39.310

Alignment\_block:  
US-09-627-165B-15 x Q96235                      ..

Align seg 1/1 to: Q96235 from: 1 to: 251





```

1 TACACGATCTGGAGCGATACCGCGT...CATAGGACAGATCCCTCT 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
141 TyraSerLeuGlnGlnGlnAlaValAspArgGlnSerTyrProle 157
48 GGGTATAGAGCAATCATCCGCTCGCGCGCT...CGTTATCCAG 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
157 uGlyProAsnSerLeuAlaGlnAlaIleSerSerLeuSerArgTyrSerG 174
95 GGGGAGCAACCGGGCCCAAGCTCGTCCCTATATATCTCATTCAGAT 144
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
174 LyGlySprAlaLysSerLeuAlaLysAlaLeuLeuValValIleGlnMet 190
145 ATCTCCGAGCGCCGAGATTCATATCCATCTTTGAGGCGTCCCAATA 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
191 ValSerGlnAlaAlaArgPheArgTyrIleGlnLeuArgIleTrpPhe 207
195 CATTAACAGCGGAGATCTTCCGACATGTACATGTCAGTCGAGCTGG 244
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
207 ValThrAspAlaIleAspPheThrProAspProLeuMetLeuSerMetG 224
245 AGACTAGTGGGGCCACATCCAGACGACGAGTCCAGAGCTTACG...GAT 291
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
224 LysLysLysTrpSerTyrMetSerLysGluIleGlnGlnAlaThrValGly 240
292 GGGCTTTTATATACCATTCGCTGGGTATATCCACCGGTAATCTTCT 341
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 GlyThrPheAlaIleValAlaGlnLeuGlnAspGlnGlyAsnAsnProI 257
342 GAGCTTGAGCAATGTTCGCGAGCTG...ATCGCCAGCTTACGACATCA 385
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
257 eAsnValThrAspPheArgArgLeuPheGlnLeuThrTyrIleAlaVal 274
386 TGTTGTTTATGTAGGAGCGACCATCT...TCCCTCCGAGCTG 426
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
274 euleuTyrGlyCysAlaProValThrSerThrSerTyrSerAsnAla 290
427 CCGCATTTGGCCGCTGTCATACGACCGCTCTGGAAATATGCGCGCT 476
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
291 IleAlaGlnIleIleArgMetProValPheArg...ValGln 304
477 CGACGATGTTACCTGCATCTTCGGAACCCAGCGTCGCGATC 519
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 yAspAspGlnLysCysThrValValAspValThrArgArgIle 318
seq_name: sp_plant:Q9M6E9
seq_documentation_block:
ID Q9M6E9 PRELIMINARY; PRI: 547 AA.
AC Q9M6E9:
DN 01-OCT-2000 (TREMUREL.15, Created)
DN 01-OCT-2000 (TREMUREL.15, Last sequence update)
DN 01-DEC-2001 (TREMUREL.19, Last annotation update)
DE RNA - GLYCOSIDASE (EC 3.2.2.22).
GN AAG.
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20102702; PubMed=10636890;
RA Lin J.I.; Tsai C.C.; Lin S.C.; Wang L.I.; Hsu C.I.; Hwang M.J.;
RT "Primary Structure and Function Analysis of the Abrus precatorius
RT Agglutinin A Chain by Site-directed Mutagenesis: ProI99 of Amphiphilic
RT alpha-Helix A Impairs Protein Synthesis Inhibitory Activity."
RT J. Biol. Chem. 275:1897-1901(2000).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
DR EMBL: AF190173; AAF28309.1; -

```

```

DR HSSP: P11140; IABR.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS0231; RICIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; UNKNOWN_1.
KW Hydroxylase; Toxin.
KW SEQUENCE 547 AA; 61248 MW; 355A325C2354A1BD CRC64;

```

```

alignment_scores:
Quality: 222.00 Length: 185
Ratio: 1.965 Gaps: 5
Percent Similarity: 61.081 Percent Identity: 34.595

```

```

alignment_block:
US-09-627-165b-15 x Q9M6E9

```

```

Align seg 1/1 to: Q9M6E9 from: 1 to: 547

```

```

1 TACACGATCTGGAGCGATACCGCGTCTAT...AGGACAGATCCCTCT 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 TyraSprLeuGlnLysTrpAlaIleGlnSerArgGlnIleSerIle 148
48 GGGTATAGAGCAATCATCCGCTCGCGCGCTCGTATCCAGCG 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 uGlyLeuGlnAlaLeuArgGlnGlyIleLysPheLeuArgSerGlyAla 165
98 GCGACCGCCGCGCCAGCTCGTCCCTTATATCCATTCAGATGATC 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
165 eAspAspGlnGlnIleAlaArgThrLeuIleValIleIleMetVal 181
148 TCCGAGCGCGCGAGTTCATATCCATCTTTGGAGCGCTCCGCAATPAC 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 AlGlnAlaIleAlaArgPheArgTyrValSerLysLeuValIleSerIle 198
198 TAAACCGGGGAGTCATTTCTCCGACATGTACATGCTCGAGCTGAGA 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
198 uSerAsnArgAlaIleAlaPheGlnProAspProSerMetLeuSerLeuGln 215
248 CTAGTTGGGGCCACATCCAGCGACGTCAGCAGCTTCAGAGGCTT 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
215 snThrTrpGluProLeuSerArgValGlnIleThrValGlnAspThr 231
298 TTTATATACCATTTCCGCTTGGGTATATCCACCGGTAATCTTCGAGCT 347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 PheProGlnAsn...ValThrIle 238
348 GAGCAATGTTCGCGAC...G 364
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
238 uIleAsnValArgGlnGlnArgValValSerSerLeuSerHisProS 255
365 TGATCCCGCACTTACGATCATCTTTGTATGTAGGACGACGACCATCT 414
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
255 eValSerAlaLeuAlaLeuMetLeuPheValCysAsnProLeuAsnAla 271
415 TCCCTCGAGCGTCCGCTATGTCGCTGATCAGACCGCGCTTGAGAAA 464
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
272 ThrGlnSer...ProLeuLeuIleArgSerValValGlnGln 284
465 TACGGCGCGCGTCGACGATGTACCTGATCTTCGACACCGCGTGC 514
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
284 nSer...LysIleCysSerSerHisTyrGlnProThrValA 297
515 GCATC 519
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 rgIle 298

```





```

seq_documentation_block:
; Sequence 16, Application US/09627165B
; GENERAL INFORMATION:
; APPLICANT: KIM, Jong-Bae
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM album coloratum, AND PROTEINS
; FILE REFERENCE: AND LECTINS ISOLATED THEREFROM
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 16
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Viscum album coloratum
; FEATURE:
; NAME/KEY: misc_feature
US-09-627-165B-16

alignment_scores:
Quality: 408.50      Length: 122
Ratio: 4.045         Gaps: 1
Percent Similarity: 82.787      Percent Identity: 69.672

alignment_block:
US-09-627-165B-13 x US-09-627-165B-16 ..
Align seg 1/1 to: US-09-627-165B-16 from: 1 to: 174

1 GCCAGATTCAATCCCATGNTGTGAGCGCTTGCAGCAATTAACAGTGG 50
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
53 AIAAGPHEANPROILEUPHETPRAGIALARGLINTRYLLEANSEREL 69
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
69 YLUGLINSERTHRLGVALGINSERTTHRASPGLYVALPHEANASN 102
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
101 GTGCAATTAAGTTCAGATTTCGCGCGTACCTTGTGACGNTGACAAATG 150
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
86 YLUGLINSERTHRLGVALGINSERTTHRASPGLYVALPHEANASN 102
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
151 CAATTAAGTTCAGATTTCGCGCGTACCTTGTGACGNTGACAAATG 200
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
103 PROPHETAGLEUAGLYLESETHRGLYANPHEVALTHRLSEUERNVA 119
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 TGGCAGCTGATCTCCAGCTTGGCGATGATGTTGTAATGCAAGTGGTC 250
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
119 IARGASPVALLLEALSERLEUVALLEMETLEUPHEVALCYARGSPRA 136
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
251 GGCGAATTCCTCTCTGACCCCTTGCAGCGCTGCTCTAGAGTCGGTC 300
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
136 TGPRESERSETSPVALARGTYTRPPROLEUVALILEATGPROVAL 152
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
301 GTGGAT.....GGCGCAACGATGTCACCTGACCTTTTCGAAC 341
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
153 LEUGUANSESGLYALAVALASPASVALTHRCYSYTHLASERGLUPR 169
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
342 CACCGTGGCATCGTA 357
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
169 OTHVALARGILEVAL 174

seq_name: /cgn2_6/prodata/2/paa/US093_COMB.pep:US-09-347-064-2

seq_documentation_block:
; Sequence 2, Application US/09347064
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE REFERENCE: 09282-5

```

```

; CURRENT APPLICATION NUMBER: US/09/347,064
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
US-09-347-064-2

alignment_scores:
Quality: 311.00      Length: 85
Ratio: 4.380         Gaps: 0
Percent Similarity: 83.529      Percent Identity: 71.765

alignment_block:
US-09-627-165B-13 x US-09-347-064-2 ..
Align seg 1/1 to: US-09-347-064-2 from: 1 to: 252

1 GCCAGATTCAATCCCATGNTGTGAGCGCTTGCAGCAATTAACAGTGG 50
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
168 AIAAGPHEANPROILEUPHETPRAGIALARGLINTRYLLEANSEREL 184
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
51 GSAGTCNTCTCCACCAACATGATGATCTGAGCTGAGACGAGTTGGC 100
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
184 YALASERHEUPROSPVALIYMETLEUGLIEGLUTHTHSETTRPG 201
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
101 GTGCAATTAAGTTCAGATTTCGCGCGTACCTTGTGACGNTGACAAATG 150
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 YLUGLINSERTHRLGVALGINSERTTHRASPGLYVALPHEANASN 217
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
151 CAATTAAGTTCAGATTTCGCGCGTACCTTGTGACGNTGACAAATG 200
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
218 PROLEUAGLEUVALIETPPROGLYANPHEVALTHRLSEUERNVA 234
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 TGGCAGCTGATCTCCAGCTTGGCGATGATGTTGTAATGCAAGTGGTC 250
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
234 IARGASPVALLLEALSERLEUVALLEMETLEUPHEVALCYSGLYLVA 251
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
251 GGCGA 255
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
251 TGP 252

seq_name: /cgn2_6/prodata/2/paa/US093_COMB.pep:US-09-347-064-8

seq_documentation_block:
; Sequence 8, Application US/09347064
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album

```

Wed Jul 31 08:46:38 2002

us-09-627-165b-13.ram

page 3

US-09-347-064-8

alignment\_scores:                      length:  
                    quality: 311.00                      85  
                    ratio: 4.380                      0  
Percent Similarity: 83.529              Percent Identity: 71.765

alignment\_block:

US-09-627-165B-13 x US-09-347-064-8 ..

Align seg 1/1 to: US-09-347-064-8 from: 1 to: 252

```
1 GCCAGATTCATCCATCCATGTCGAGAGCTTCGCCGCAATTAACAGTGG 50
|||||
167 ALaarGpHeasnProlleuTriPaRgAlaRgSlTyRtlleasnsErcI 183
51 GGAGTCNTCTCCACCAACATGPACTGCTGAGCTGGAGAGAGTGGG 100
| ||| |||
183 yLaseRpheuPrhoSpvaLyRmetleuGlueuGluThrSeRtrpG 200
101 GTGGCAATCCACCCAGTCGAGCTTGCGAGATCGAAGATGCATTTTAATAC 150
|||||
200 lYngInSertThrgInValGlnHIsSerThrsPglYalPheasnsAn 216
151 CAATAGATTCAGATTCCGCCGCTAACCTTGTGACGNTGACCAATGT 200
|||||
217 ProIlleargLeuAlaIleProProGlyAsnPheValThrlleuThrsAnVa 233
201 TCGGAGATTCAGATTCGAGCTTGCGAGATCATGTTGTCGATGCAAGTGGTC 250
|||||
233 lArGaSpvaIlleAlaseRleuAlaIleMetleuPheValCysGlyGluA 250
251 GGCCA 255
|||||
250 rGPro 251
```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-347-064-2

seq\_documentation\_block:

```
; Sequence 2, Application US/09347064A
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
US-09-347-064-2
```

alignment\_scores:                      length:  
                    quality: 311.00                      85  
                    ratio: 4.380                      0  
Percent Similarity: 83.529              Percent Identity: 71.765

alignment\_block:

US-09-627-165B-13 x US-09-347-064-2 ..

Align seg 1/1 to: US-09-347-064-2 from: 1 to: 252

```
1 GCCAGATTCATCCATCCATGTCGAGAGCTTCGCCGCAATTAACAGTGG 50
|||||
168 ALaarGpHeasnProlleuTriPaRgAlaRgSlTyRtlleasnsErcI 184
51 GGAGTCNTCTCCACCAACATGPACTGCTGAGCTGGAGAGAGTGGG 100
| ||| |||
184 yLaseRpheuPrhoSpvaLyRmetleuGlueuGluThrSeRtrpG 201
101 GTGGCAATCCACCCAGTCGAGCTTGCGAGATCGAAGATGCATTTTAATAC 150
|||||
201 lYngInSertThrgInValGlnHIsSerThrsPglYalPheasnsAn 217
151 CAATAGATTCAGATTCCGCCGCTAACCTTGTGACGNTGACCAATGT 200
|||||
218 ProIlleargLeuAlaIleProProGlyAsnPheValThrlleuThrsAnVa 234
201 TCGGAGATTCAGATTCGAGCTTGCGAGATCATGTTGTCGATGCAAGTGGTC 250
|||||
234 lArGaSpvaIlleAlaseRleuAlaIleMetleuPheValCysGlyGluA 251
251 GGCCA 255
|||||
251 rGPro 252
```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-347-064-8

seq\_documentation\_block:

```
; Sequence 8, Application US/09347064A
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
US-09-347-064-8
```

alignment\_scores:                      length:  
                    quality: 311.00                      85  
                    ratio: 4.380                      0  
Percent Similarity: 83.529              Percent Identity: 71.765

alignment\_block:

US-09-627-165B-13 x US-09-347-064-8 ..

Align seg 1/1 to: US-09-347-064-8 from: 1 to: 252

```
1 GCCAGATTCATCCATCCATGTCGAGAGCTTCGCCGCAATTAACAGTGG 50
|||||
167 ALaarGpHeasnProlleuTriPaRgAlaRgSlTyRtlleasnsErcI 183
51 GGAGTCNTCTCCACCAACATGPACTGCTGAGCTGGAGAGAGTGGG 100
| ||| |||
183 yLaseRpheuPrhoSpvaLyRmetleuGlueuGluThrSeRtrpG 200
101 GTGGCAATCCACCCAGTCGAGCTTGCGAGATCGAAGATGCATTTTAATAC 150
|||||
200 lYngInSertThrgInValGlnHIsSerThrsPglYalPheasnsAn 216
```

```
151 CAATAGATTCAGATTCCCGGCTGTAAGTGTGAGTCAGTACGATG 200
|||||  |||  |||||  |||||  |||||  |||||  |||||  |||||
217 ProileargleuAlaIleProProGlyAsnPhenValThrIleuThrAsnVa 233
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 TCCGAGCTGATCTCCAGTTCGGGATCATGTTGTTCGATGACATGCTG 250
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
233 IatgAspValIleAlaSerIleuAlaIleMetIleuPheValCysIleuA 250
251 GGCGA 255
|||||
250 r9pPro 251

seq_name: /cgn2_6/prodata/2/paa/US096_COMB.pep:US-09-627-165b-2
seq_documentation_block:
; Sequence 2, Application US/09627165B
; GENERAL INFORMATION:
; APPLICANT: KIM, Jong-Bae
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM ALBUM COLORATUM, AND PROTEINS
; FILE REFERENCE: Korean Mistletoe Lectin
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 2
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Viscum album coloratum
; NAME/KEY: misc-feature
US-09-627-165b-2

alignment_scores:
Quality: 303.00 Length: 88
Ratio: 4.151 Gaps: 0
Percent Similarity: 82.955 Percent Identity: 69.318

alignment_block:
US-09-627-165b-13 x US-09-627-165b-2 ..
Align seg 1/1 to: US-09-627-165b-2 from: 1 to: 254

1 GCCAGATTCAATCCCATCNCNTGTGAGGCTTCGCCGCAAAATTACAGTGG 50
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
167 AlaArgPheAsnProIleuThrPArgAlaArgGlnTyrIleSerSerGI 183
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
51 GGAGTCNCTCTCCACCAACATGTACATGCTCGAGCTGAGAGAGAGTGGG 100
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
183 YglYserPheIeuProAspThrTyrIleuGlnIleuGlnThrSerTrp 200
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
101 GTGACAAATCCACCAAGTCCAGAGTCGACAGATGCGATTTTTAATAC 150
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
200 YglGlnIleSerThrGlnValGlnIleSerThrAspGlyValIlePheAsnA 216
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
151 CAATAGATTCAGATTCCCGGCTGTAAGTGTGAGTCAGTACGATG 200
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
217 ProileargleuThrIleSerThrGlyValIlePheValThrIleuSerAsnVa 233
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 TCCGAGCTGATCTCCAGTTCGGGATCATGTTGTTCGATGACATGCTG 250
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
233 IatgAspValIleAlaSerIleuAlaIleMetIleuPheValCysIleuA 250
251 GGCGATTCTCCCT 264
|||||  |||||
250 r9pProSerSer 254

seq_name: /cgn2_6/prodata/2/paa/US096_COMB.pep:US-09-627-165b-4
seq_documentation_block:
; Sequence 4, Application US/09627165B
; GENERAL INFORMATION:
; APPLICANT: KIM, Jong-Bae
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM ALBUM COLORATUM, AND PROTEINS
; FILE REFERENCE: Korean Mistletoe Lectin
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Viscum album coloratum
; NAME/KEY: misc-feature
US-09-627-165b-6

APPLICANT: KIM, Jong-Bae
TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM ALBUM COLORATUM, AND PROTEINS
FILE REFERENCE: Korean Mistletoe Lectin
CURRENT APPLICATION NUMBER: US/09/627,165B
CURRENT FILING DATE: 2000-07-27
NUMBER OF SEQ ID NOS: 16
SEQ ID NO 4
LENGTH: 254
TYPE: PRT
ORGANISM: Viscum album coloratum
FEATURE:
NAME/KEY: misc-feature
LOCATION: 240
OTHER INFORMATION: Xaa = any amino acid
US-09-627-165b-4

alignment_scores:
Quality: 298.00 Length: 88
Ratio: 4.139 Gaps: 0
Percent Similarity: 81.818 Percent Identity: 68.182

alignment_block:
US-09-627-165b-13 x US-09-627-165b-4 ..
Align seg 1/1 to: US-09-627-165b-4 from: 1 to: 254

1 GCCAGATTCAATCCCATCNCNTGTGAGGCTTCGCCGCAAAATTACAGTGG 50
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
167 AlaArgPheAsnProIleuThrPArgAlaArgGlnTyrIleSerSerGI 183
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
51 GGAGTCNCTCTCCACCAACATGTACATGCTCGAGCTGAGAGAGAGTGGG 100
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
183 YglYserPheIeuProAspThrTyrIleuGlnIleuGlnThrSerTrp 200
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
101 GTGACAAATCCACCAAGTCCAGAGTCGACAGATGCGATTTTTAATAC 150
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
200 YglGlnIleSerThrGlnValGlnIleSerThrAspGlyValIlePheAsnA 216
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
151 CAATAGATTCAGATTCCCGGCTGTAAGTGTGAGTCAGTACGATG 200
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
217 ProileargleuThrIleSerThrGlyValIlePheValThrIleuSerAsnVa 233
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 TCCGAGCTGATCTCCAGTTCGGGATCATGTTGTTCGATGACATGCTG 250
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
233 IatgAspValIleAlaSer***AlaIleMetIleuPheValCysIleuA 250
251 GGCGATTCTCCCT 264
|||||  |||||
250 r9pProSerSer 254

seq_name: /cgn2_6/prodata/2/paa/US096_COMB.pep:US-09-627-165b-6
seq_documentation_block:
; Sequence 6, Application US/09627165B
; GENERAL INFORMATION:
; APPLICANT: KIM, Jong-Bae
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM ALBUM COLORATUM, AND PROTEINS
; FILE REFERENCE: Korean Mistletoe Lectin
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Viscum album coloratum
; NAME/KEY: misc-feature
US-09-627-165b-6
```

alignment\_scores:                    length: 87  
                  quality: 298.00  
                  ratio: 4.197  
Percent similarity: 81.609      Percent identity: 68.966

alignment\_block:  
US-09-627-165b-13 x US-09-627-165b-6 ..

Align seg 1/1    to: US-09-627-165b-6    from: 1    to: 256

```
1 GCCAGATTCAATCCATCCTGTCGAGAGCTTCGCGGCAAAATTACAGTGG 50
|||||
169 ALaargpheaSnprolleuTrparGaLaargInTyrlleaSnserG1 185
51 GGAGTCTCTCCACCAACATGATGCTGAGCTGAGAGAGAGTGGG 100
|||||
185 yvalserTyrlleuproaSpvalTyrlmetleuGluLeuGluAlaserTrpg 202
101 GTGCAATCCACCCCAAGTCCAGACAGTCCAGAGATGTTTAAATAC 150
|||||
202 lYgInGInserThInGInValGInserThraSpelyAlaPheasnAn 218
151 CAATTAAGATTGACATTTCCGCGGTAACTTTGTGACGNTGAGCAATG 200
|||||
219 ProlleargleuGlylleSerThInGlyAsnPhaveAlTrpleuSerAsnva 235
201 TCGCAGAGTATCTCCAGCTTGGCGATCATGTTGTCGAATGAGAGTGC 250
|||||
225 lArGAspValleAlaserleuGlylleMetValPheValCyArGAspA 252
251 GGCCATTCCTCC 261
252 rGserSerSer 255
```

seq\_name: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:US-08-101-329-6

seq\_documentation\_block:  
Sequence 6, Application US/08101329

GENERAL INFORMATION:  
APPLICANT: Wood, Mark S.  
APPLICANT: Gould, Robert M.  
APPLICANT: Kelleher, Peter J.  
APPLICANT: Wallace, Thomas L.  
TITLE OF INVENTION: SINGLE CHAIN-IMMUNOTOXIN  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PREVENTING SECONDARY  
TITLE OF INVENTION: CATARACTS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish and Richardson  
STREET: 2200 Sand Hill Road, Suite 100  
CITY: Menlo Park  
STATE: California  
COUNTRY: US  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/101,329  
FILING DATE: 03-AUG-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rae-Venter, Barbara  
REGISTRATION NUMBER: 32,750  
REFERENCE/DOCKET NUMBER: 05936/021001  
TELEPHONE: (415) 854-3277  
TELEFAX: (415) 854-0875  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 554 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-101-329-6

alignment\_scores:                    length: 94  
                  quality: 140.50  
                  ratio: 2.266  
Percent similarity: 65.957      Percent identity: 35.106

alignment\_block:  
US-09-627-165b-13 x US-08-101-329-6 ..

Align seg 1/1    to: US-08-101-329-6    from: 1    to: 554

```
1 GCCAGATTCAATCCATCCTGTCGAGAGCTTCGCGGCAAAATTACAGTGG 50
|||||
472 ALaargpheaSnprolleuTrparGaLaargInTyrlleaSnserG1 472
51 GGAGTCTCTCCACCAACATGATGCTGAGCTGAGAGAGAGTGGG 100
|||||
472 nArGAspValleAlaserleuGlylleMetValPheValCyArGAspA 489
101 GTGCAATCCACCCCAAGTCCAGACAGTCCAGAGATGTTTAAATAC 150
|||||
489 lYgInGInserThInGInValGInserThraSpelyAlaPheasnAn 505
151 CAATTAAGATTGACATTTCCGCGGTAACTTTGTGACGNTGAGCAATG 200
|||||
506 ProlleargleuGlylleSerThInGlyAsnPhaveAlTrpleuSerAsnva 522
201 TCGCAGAGTATCTCCAGCTTGGCGATCATGTTGTCGAATGAGAGTGC 250
|||||
522 lserlleuAlaserleuGlylleMetValPheValCyArGAspA 539
251 GGCCA.....TTCTCCTCTCTGACAC 273
539 rGserSerSerGlnPheGlySerArGSerHis 549
```

seq\_name: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:US-07-787-567A-1

seq\_documentation\_block:  
Sequence 1, Application US/07787567A

GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Steve F.  
APPLICANT: Lane, Julie A.  
TITLE OF INVENTION: Materials Comprising and Methods of  
TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/787,567A  
FILING DATE: 19911104  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Gruber, Lewis S.  
REGISTRATION NUMBER: 30,060  
REFERENCE/DOCKET NUMBER: 27129/30545

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-787-567A-1

alignment_scores:
  Quality: 138.00      Length: 87
  Ratio: 2.379         Gaps: 0
  Percent Similarity: 66.667   Percent Identity: 35.632

alignment_block:
US-09-627-165b-13 x US-07-787-567A-1 ..

Align seg 1/1 to: US-07-787-567A-1 from: 1 to: 267

1  GCCAGATTCATCCATCCATNTGTGAGGCTTCGCGCGCAATTAACAGTGC 50
  |||||.....| | | | | .....| | | | | .....| | | | | ..
179 AlAArgPheGlnTyrTlleGluGlyGluMetArgThrArgIleArgTyrAs 195
  : : : : | | | | | : : : : | | | | | : : : : | | | | |
51 GGAATCCTCTCCACCAACATGATGCTGACGCTGAGACGAGTGGG 100
  : : : : | | | | | : : : : | | | | | : : : : | | | | |
195 nArGAgSerAlaProAspProSerValIleThrLeuGluAsnSerTrpG 212
  : : : : | | | | | : : : : | | | | | : : : : | | | | |
101 GTGACATTCACCCCAAGTCAGACGATCCAAAGATGCGATTTTAATACC 150
  ||||| | | | | | : : | | | | | : : | | | | | | | | | : :
212 LyArgLeuSerThrAlaIleGlnIleuSerAsnGlnGlyAlaPheAlaSer 228
  ||||| | | | | | : : | | | | | : : | | | | | | | | | : :
229 ProIleGlnLeuGlnArgArgAsnGlySerLysPheSerValTyrAspVa 245
  : : : : | | | | | : : : : | | | | | : : : : | | | | |
201 TCAGCAGCTGATCTCCAGCTTGCGGATGATGTTGTCGAATGACAGTGTC 250
  | : : : | | | | | : : | | | | | : : | | | | | | | | |
245 lSerIleLeuIleProIleIleAlaLeuMetValTyrArgCysAlaProp 262
  251 GGCCATTCTCC 261
  || | | |
262 roProSerSer 265

seq_name: /cgn2_6/ptodata/2/paa/US084_COMB.pep:US-08-406-801-34

seq_documentation_block:
; Sequence 34, Application US/08406801
; GENERAL INFORMATION:
; APPLICANT: Wright, Andrew F.
; APPLICANT: Blakey, David C.
; APPLICANT: Flitton, John E.
; APPLICANT: Lindholm, Delit
; APPLICANT: Lind, Peter
; APPLICANT: Holmgren, Jan
; TITLE OF INVENTION: Conjugates
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman, Darby & Cushman
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/406,801
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,269
; FILING DATE: 02-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Kokulis, Paul N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: PNK/3893/96491/MJM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
;
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-406-801-34

alignment_scores:
  Quality: 138.00      Length: 87
  Ratio: 2.379         Gaps: 0
  Percent Similarity: 66.667   Percent Identity: 35.632

alignment_block:
US-09-627-165b-13 x US-08-406-801-34 ..

Align seg 1/1 to: US-08-406-801-34 from: 1 to: 267

1  GCCAGATTCATCCATCCATNTGTGAGGCTTCGCGCGCAATTAACAGTGG 50
  |||||.....| | | | | .....| | | | | .....| | | | | ..
179 AlAArgPheGlnTyrTlleGluGlyGluMetArgThrArgIleArgTyrAs 195
  : : : : | | | | | : : : : | | | | | : : : : | | | | |
51 GGAATCCTCTCCACCAACATGATGCTGACGCTGAGACGAGTGGG 100
  : : : : | | | | | : : : : | | | | | : : : : | | | | |
195 nArGAgSerAlaProAspProSerValIleThrLeuGluAsnSerTrpG 212
  : : : : | | | | | : : : : | | | | | : : : : | | | | |
101 GTGACATTCACCCCAAGTCAGACGATCCAAAGATGCGATTTTAATACC 150
  ||||| | | | | | : : | | | | | : : | | | | | | | | | : :
212 LyArgLeuSerThrAlaIleGlnIleuSerAsnGlnGlyAlaPheAlaSer 228
  ||||| | | | | | : : | | | | | : : | | | | | | | | | : :
229 ProIleGlnLeuGlnArgArgAsnGlySerLysPheSerValTyrAspVa 245
  : : : : | | | | | : : : : | | | | | : : : : | | | | |
151 CAATTAAGATTCAGATTTCGCGCGTAACTTTGTGACAGNTGACCAATGT 200
  ||||| | | | | | : : | | | | | : : | | | | | | | | | : :
201 TCAGCAGCTGATCTCCAGCTTGCGGATGATGTTGTCGAATGACAGTGTC 250
  | : : : | | | | | : : | | | | | : : | | | | | | | | |
245 lSerIleLeuIleProIleIleAlaLeuMetValTyrArgCysAlaProp 262
  251 GGCCATTCTCC 261
  || | | |
262 roProSerSer 265

seq_name: /cgn2_6/ptodata/2/paa/US095_COMB.pep:US-09-538-873-1

seq_documentation_block:
; Sequence 1, Application US/09538873
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETTE, VICTOR F.
; APPLICANT: SWALSHAW, JOAN
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
; FILE REFERENCE: US/09/538,873
; CURRENT APPLICATION NUMBER: US/09/538,873
; CURRENT FILING DATE: 2000-03-30
; EARLIER APPLICATION NUMBER: 60/126,826
; EARLIER FILING DATE: 1999-03-30
```

```

PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/09/136,389
  FILING DATE: 18-AUG-1998
  APPLICATION NUMBER: 08/646,360
  FILING DATE: 13-MAY-1996
  APPLICATION NUMBER: PCT/US94/05348
  FILING DATE: 12-MAY-1994
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/054,691
    FILING DATE: 12-MAY-1993
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 07/988,430
      FILING DATE: 09-DEC-1992
      PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 07/901,707
        FILING DATE: 19-JUN-1992
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: US 07/787,567
          FILING DATE: 04-NOV-1991
          ATTORNEY/AGENT INFORMATION:
            NAME: MONICHOIAS, Janet M.
            REGISTRATION NUMBER: 32,918
            REFERENCE/DOCKET NUMBER: 200-70.P4
            TELECOMMUNICATION INFORMATION:
              TELEPHONE: 312/707-8889
              TELEFAX: 312/707-9155
              TELEX: 650 388-1248
          INFORMATION FOR SEQ ID NO: 1:
            SEQUENCE CHARACTERISTICS:
              LENGTH: 267 amino acids
              TYPE: amino acid
              TOPOLOGY: linear
            MOLECULE TYPE: protein
US-09-610-838-1

alignment_scores:
  Quality: 138.00      Length: 87
  Ratio: 2.379        Gaps: 0
  Percent Similarity: 66.667      Percent Identity: 35.632

alignment_block:
US-09-627-1658-13 x US-09-610-838-1  ..

Align seg 1/1  to: US-09-610-838-1  from: 1  to: 267

1  GCCGACATTCATCCATCCTGTCGAGGCTTCCGCCGCAATTACACGTGG  50
   |||||:::  ||  ::::||  ::::||  ::::||  :::
179  AlarqphedlnyrlieglnylsumetargthrArglllearyTyrAs  195
   51  GGAGCTCTCCCAACCAATGTAATGCTGAGCTGAGACGAGTGGG  100
      ::::  ::::  ::::  ::::  ::::  ::::  ::::
195  naryargSerlalarphosproSerValllerThrleuglunansertIPG  212
   101  GTCGACATTCACCCCAAGTCCACGACATGCCAAGATGGCAATTTAATAC  150
      ||||  ||||  ::||::||::||::||  ||  :::
212  lyarglueserThrallelglnlserasnlglnlalaIlealaser  228
   151  CAATAGATGGAGATTTCGCCGCGGTAACTTTGTGACGTGACATGT  200
      ||:::|||||  ||:::  :::  :::
229  ProileglnlunlunargatargasnclSerlyPheSerVallyrAspya  245
   201  TCGGACGCTGATCTCCAGCTTGGCGATCATGTTCTTCATGAGAGTGC  250
      |  ::||  ::||::||::||::||::||::||::||::||::||::||
245  lserleuleulleprollerlealeuMetVallyrArgcyahalarop  262
   251  GGCCATTCCTCC  261
      |||  |||
262  roPserSer  265

seq_name: /cgn2_6/ptodata/2/paa/JS096_COMB_pep.us-09-668-419-1
seq_documentation_block:

```

Wed Jul 31 08:46:38 2002

us-09-627-165b-13.rapm

**Page 8**

```

Sequence 1, Application US/09668419
GENERAL INFORMATION:
APPLICANT: VITETTA, ELLEN S.
APPLICANT: GHETIE, VICTOR F.
APPLICANT: SMALLSHAW, JOAN E.
APPLICANT: BALUNA, ROXANA G.
TITLE OF INVENTION: RICIN A CHAIN MUTANTS LACKING ENZYMATIC ACTIVITY AS VACCINES TO
TITLE OF INVENTION: AGAINST AEROSOLIZED RICIN
FILE REFERENCE: UDS:723US
CURRENT APPLICATION NUMBER: US/09/666, 419
CURRENT FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 267
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-666-419-1

```

```
alignment_scores:
  Quality: 138.00
  Ratio: 2.379
  Percent Similarity: 66.667
  Length: 87
  Gaps: 0
  Percent Identity: 35.632
```

```
alignment_block: 118-09-6668-419-1
```

Align seg 1/1 to: US-09-668-419-1 from: 1 to: 267

[illegible]





```

281 g11eVal 283
seq_name: /sgn2_6/prodata/1/Paa/US09_NEW_COMB pep:us-09-601-667B-1
seq_documentation_block:
  Sequence 1, Application US/09601667B
  GENERAL INFORMATION:
    APPLICANT: Morris, Peter
    APPLICANT: Stiefel, Thomas
    APPLICANT: Voelter, Wolfgang
    APPLICANT: Welters, Peter
    TITLE OF INVENTION: Recombinant Mistletoe Lectins
    FILE REFERENCE: 29841/36636
    CURRENT APPLICATION NUMBER: US/09/601.667B
    PRIOR FILING DATE: 2000-10-06
    PRIOR APPLICATION NUMBER: PCT/EP99/00696
    PRIOR FILING DATE: 1999-02-03
    PRIOR APPLICATION NUMBER: D 199 04 210.8
    NUMBER OF SEQ ID NOS: 41
    SEQ ID NO 1
    LENGTH: 533
    TYPE: PRT
    ORGANISM: Artificial Sequence
  FEATURE:
    OTHER INFORMATION: mistletoe lectin
  NAME/KEY: SITE
  LOCATION: 15
  OTHER INFORMATION: product="Xaa is Asp or Glu"
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 63
    OTHER INFORMATION: product="Xaa is Gly or Gln"
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 66
    OTHER INFORMATION: product="Xaa is Ile or Val"
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 75
    OTHER INFORMATION: product="Xaa is Leu or Ala"
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 107
    OTHER INFORMATION: product="Xaa is missing"
  OTHER INFORMATION: /label= Xaa5
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 113
    OTHER INFORMATION: product="Xaa is Asn or Thr"
  OTHER INFORMATION: /label= Xaa6
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 117
    OTHER INFORMATION: product="Xaa is Pro or Thr"
  OTHER INFORMATION: /label= Xaa7
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 134
    OTHER INFORMATION: product="Xaa is Asp or Glu"
  OTHER INFORMATION: /label= Xaa8
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 141
    OTHER INFORMATION: product="Xaa is Ser or Thr"
  OTHER INFORMATION: /label= Xaa9
  NAME/KEY: SITE
  LOCATION: 145
  OTHER INFORMATION: product="Xaa is Phe or Tyr"
  OTHER INFORMATION: /label= Xaa10
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 152
    OTHER INFORMATION: product="Xaa is Thr or Ala"
  OTHER INFORMATION: /label= Xaa11
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 177
    OTHER INFORMATION: product="Xaa is Ala or Tyr"
  OTHER INFORMATION: /label= Xaa12
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 180
    OTHER INFORMATION: product="Xaa is Tyr or Asp"
  OTHER INFORMATION: /label= Xaa13
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 185
    OTHER INFORMATION: product="Xaa is Ala or Glu"
  OTHER INFORMATION: /label= Xaa14
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 191
    OTHER INFORMATION: product="Xaa is Val or Met"
  OTHER INFORMATION: /label= Xaa15
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 210
    OTHER INFORMATION: product="Xaa is Ile or Phe"
  OTHER INFORMATION: /label= Xaa16
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 224
    OTHER INFORMATION: product="Xaa is Pro or Ser"
  OTHER INFORMATION: /label= Xaa17
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 225
    OTHER INFORMATION: product="Xaa is Pro or Thr"
  OTHER INFORMATION: /label= Xaa18
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 232
    OTHER INFORMATION: product="Xaa is Thr or Ser"
  OTHER INFORMATION: /label= Xaa19
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 236
    OTHER INFORMATION: product="Xaa is Asp or Ser"
  OTHER INFORMATION: /label= Xaa20
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 287
    OTHER INFORMATION: product="Xaa is Asn or Ser"
  OTHER INFORMATION: /label= Xaa21
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 290
    OTHER INFORMATION: product="Xaa is Cys or Arg"
  OTHER INFORMATION: /label= Xaa22
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 325
    OTHER INFORMATION: product="Xaa is Gly or Asn"
  OTHER INFORMATION: /label= Xaa23
  FEATURE:
    NAME/KEY: SITE
    LOCATION: 364
    OTHER INFORMATION: product="Xaa is Gly or Asp"
  OTHER INFORMATION: /label= Xaa24

```

```

FEATURE:
NAME/KEY: SITE
LOCATION: 426
OTHER INFORMATION: product= "Xaa Is Gly or Gln"
OTHER INFORMATION: /label= Xaa25
FEATURE:
NAME/KEY: SITE
LOCATION: 435
OTHER INFORMATION: product= "Xaa Is Val or Asp"
OTHER INFORMATION: /label= Xaa26
FEATURE:
NAME/KEY: SITE
LOCATION: 439
OTHER INFORMATION: product= "Xaa Is Gln or Lys"
OTHER INFORMATION: /label= Xaa27
FEATURE:
NAME/KEY: SITE
LOCATION: 442
OTHER INFORMATION: product= "Xaa Is Gly or missing"
OTHER INFORMATION: /label= Xaa28
FEATURE:
NAME/KEY: SITE
LOCATION: 443
OTHER INFORMATION: product= "Xaa Is Arg or Lys"
OTHER INFORMATION: /label= Xaa29
FEATURE:
NAME/KEY: SITE
LOCATION: 464
OTHER INFORMATION: product= "Xaa Is Cys or Ser or Val"
OTHER INFORMATION: /label= Xaa30
FEATURE:
NAME/KEY: SITE
LOCATION: 480
OTHER INFORMATION: product= "Xaa Is Ala or Gly"
OTHER INFORMATION: /label= Xaa
FEATURE:
NAME/KEY: SITE
LOCATION: 481
OTHER INFORMATION: product= "Xaa Is Gly or Ala"
OTHER INFORMATION: /label= Xaa32
FEATURE:
NAME/KEY: SITE
LOCATION: 483
OTHER INFORMATION: product= "Xaa Is Ser or Gly"
OTHER INFORMATION: /label= Xaa33
FEATURE:
NAME/KEY: SITE
LOCATION: 484
OTHER INFORMATION: product= "Xaa Is Gly or Ser"
OTHER INFORMATION: /label= Xaa34
FEATURE:
NAME/KEY: SITE
LOCATION: 493
OTHER INFORMATION: product= "Xaa Is Gly or Tyr"
OTHER INFORMATION: /label= Xaa35
FEATURE:
NAME/KEY: SITE
LOCATION: 500
OTHER INFORMATION: product= "Xaa Is Asn or Ser or Thr or Lys"

alignment_scores:
Quality: 373.00      Length: 119
Ratio: 3.885         Gaps: 1
Percent Similarity: 80.672      Percent Identity: 65.546

alignment block:
US-09-627-165B-13 x US-09-601-667B-1 ..
Align seq 1/1 to: US-09-601-667B-1 from: 1 to: 533
1 GCCAGATTCATCCATCCTGAGAGCTTCCGCGCAATTAACAGTGG 50
|||||
168 AlaArgPheAsnProIleuLeuTyrParg**ArgGln**IleAsnSerG 184

```

```

51 GGAGTCNTCTCCACCAACATGATCATCTGAGCTTGACACAGATTGGG 100
| ||| |||||
184 y***SerPheLeuProasp**TyrMetLeuGluLeuGluThrSerTrp 201
101 GTCCGACATCCACCCAAAGTCAGCAGTCGCAAGATGCGATTATTAAC 150
|||||
201 IyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsnAsn 217
151 CAATTAAGATGCGATTCCTCCGCGGTAAGTTGTGACGNTGACCAATG 200
|||||
218 Pro***ArgLeuAlaIle*****GlyAsnPheValThrLeu**AsnVa 234
201 TCGGAGATGATCTCCAGCTTGCGCATGATGTTGCGAATGACAGTGC 250
|||||
234 Iarg**ValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGluA 251
251 GCGCATTCCTCTCTCGACACACCTTCCGCGCTGCTTGAAGTCCGTC 300
|||||
251 rProSerSerSerAspValArgTyrTrpProLeuValIleArgProVal 267
301 GTGGATGCGGCGCAAGATGTCACCTGCACTTTTCGGAACCGCGTGG 350
||| |||||
268 Ile.....AlaAspAspValThrCysSerAlaSerGluProThrValAr 282
351 CATCGTA 357
|||
282 gIleVal 284

seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-40
seq documentation block:
: Sequence 40, Application US/09601667B
: GENERAL INFORMATION:
: APPLICANT: Morris, Peter
: APPLICANT: Stiefel, Thomas
: APPLICANT: Voelter, Wolfgang
: TITLE OF INVENTION: Recombinant Mistletoe Lectins
: FILE REFERENCE: 29841/36636
: CURRENT APPLICATION NUMBER: US/09/601,667B
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: PCT/EP99/00696
: PRIOR APPLICATION NUMBER: D 198 04 210.8
: NUMBER OF SEQ ID NOS: 41
: SEQ ID NO 40
: LENGTH: 534
: TYPE: PRP
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: mistletoe lectin
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (15)..(15)
: OTHER INFORMATION: Xaa Is Asp or Glu
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (63)..(63)
: OTHER INFORMATION: Xaa Is Gly or Gln
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (66)..(66)
: OTHER INFORMATION: Xaa Is Ile or Val
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (75)..(75)
: OTHER INFORMATION: Xaa Is Leu or Ala
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (114)..(114)
: OTHER INFORMATION: Xaa Is Asn or Thr

```

FEATURE:  
NAME/KEY: SITE  
LOCATION: (118)..(118)  
OTHER INFORMATION: Xaa is Pro or Thr  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (135)..(135)  
OTHER INFORMATION: Xaa is Asp or Glu  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (142)..(142)  
OTHER INFORMATION: Xaa is Ser or Thr  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (146)..(146)  
OTHER INFORMATION: Xaa is Phe or Tyr  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (153)..(153)  
OTHER INFORMATION: Xaa is Thr or Ala  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (178)..(178)  
OTHER INFORMATION: Xaa is Ala or Tyr  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (181)..(181)  
OTHER INFORMATION: Xaa is Tyr or Asp  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (186)..(186)  
OTHER INFORMATION: Xaa is Ala or Glu  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (192)..(192)  
OTHER INFORMATION: Xaa is Val or Met  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (220)..(220)  
OTHER INFORMATION: Xaa is Ile or Phe  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (225)..(225)  
OTHER INFORMATION: Xaa is Pro or Ser  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (226)..(226)  
OTHER INFORMATION: Xaa is Pro or Thr  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (233)..(233)  
OTHER INFORMATION: Xaa is Thr or Ser  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (237)..(237)  
OTHER INFORMATION: Xaa is Asp or Ser  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (288)..(288)  
OTHER INFORMATION: Xaa is Asn or Ser  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (291)..(291)  
OTHER INFORMATION: Xaa is Cys or Arg  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (326)..(326)  
OTHER INFORMATION: Xaa is Gly or Asn  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (365)..(365)  
OTHER INFORMATION: Xaa is Gly or Asp  
FEATURE:

NAME/KEY: SITE  
LOCATION: (427)..(427)  
OTHER INFORMATION: Xaa is Gly or Gln  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (436)..(436)  
OTHER INFORMATION: Xaa is Val or Asp  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (440)..(440)  
OTHER INFORMATION: Xaa is Gln or Lys  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (443)..(443)  
OTHER INFORMATION: Xaa is Gly or missing  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (444)..(444)  
OTHER INFORMATION: Xaa is Arg or Lys  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (465)..(465)  
OTHER INFORMATION: Xaa is Cys or Ser or Val  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (481)..(481)  
OTHER INFORMATION: Xaa is Ala or Gly  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (482)..(482)  
OTHER INFORMATION: Xaa is Gly or Ala  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (484)..(484)  
OTHER INFORMATION: Xaa is Ser or Gly  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (485)..(485)  
OTHER INFORMATION: Xaa is Gly or Ser  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (494)..(494)  
OTHER INFORMATION: Xaa is Gly or Tyr  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (501)..(501)  
OTHER INFORMATION: Xaa is Asn or Ser or Thr or Lys  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (502)..(502)  
OTHER INFORMATION: Xaa is Ser or Gly  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (503)..(503)  
OTHER INFORMATION: Xaa is Leu or Pro  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (504)..(504)  
OTHER INFORMATION: Xaa is Ala or Met  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (505)..(505)  
OTHER INFORMATION: Xaa is Met or Val  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (534)..(534)  
OTHER INFORMATION: Xaa is Pro or Phe  
US-09-601-667B-40

alignment\_scores:  
Quality: 373.00 Length: 119  
Ratio: 3.885 Gaps: 1

Percent Similarity: 80.672 Percent Identity: 65.546

## alignment\_block:

US-09-627-165b-13 x US-09-601-667b-40 ..

Align seg 1/1 to: US-09-601-667b-40 from: 1 to: 534

```
1 GCCAGATTCAATCCCATCNTGTGGAGGCTTCCGCCGCAATTAACAGTGG 50
|||||
169 AlArGpHeAsnPrOllLeuTrPaRg***ArgGln**IleAsnSerGI 185
51 GGAGTCNTCTCCACCAACATGTCATGCTGCAGCTGGAGACGAGTGGG 100
|||||
185 y**SerPheLeuProAsp**TyrMetLeuGlnLeuGlnThrSetrPg 202
101 GTCGACATCCACCCAGTCCAGCAGTCCAAAGATGGCATTTTAAATACC 150
|||||
202 YgInGlnSerThrGlnValGlnHisSerThrAspGlyAlaPheAsnAsn 218
151 CAAATTAAGATTGCAGATTTCGCCGGTAATTGTGTGACGNTGACAAATGT 200
|||||
219 Pro**ArgLeuAlaIle***GlyAsnPheValThrLeu**AsnVa 235
201 TCGGACGTGATCTCCAGCTTGGCGATCATGTGTTCGAATGCAGTGTGC 250
|||||
235 lArg**ValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGlnu 252
251 GGCCATTCTCCTCT 300
|||||
252 rGProSerSerSerAspValAlaGlyTrpProLeuValIleAlaGProVal 268
301 GTGGATGCGGCCAGATGTCACCTGCAGCTTTTCCGAACCCACCGTGGC 350
|||||
269 Ile.....AlaAspAspValThrCysSerAlaSerGlnurProthrValAr 283
351 CATCGTA 357
|||||
283 gIleVal 285
```

seq\_name: /cgn2\_6/prodata/1/paa/US09\_NEW\_COMB.pep:US-09-601-667b-5

## seq\_documentation\_block:

; Sequence 5, Application US/09601667B  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Peter  
; APPLICANT: Stiefel, Thomas  
; APPLICANT: Voelter, Wolfgang  
; APPLICANT: Welters, Peter  
; TITLE OF INVENTION: Recombinant Mistletoe Lectins  
; FILE REFERENCE: 29841/36636  
; CURRENT APPLICATION NUMBER: US/09/601,667B  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: PCT/EP99/00696  
; PRIOR FILING DATE: 1999-02-03  
; PRIOR APPLICATION NUMBER: D 198 04 210.8  
; NUMBER OF SEQ ID NOS: 41  
; SEQ ID NO 5  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Lectin A2  
US-09-601-667b-5

alignment\_scores:  
Quality: 327.00 Length: 88  
Ratio: 4.247 Gaps: 0  
Percent Similarity: 87.500 Percent Identity: 73.864

## alignment\_block:

US-09-627-165b-13 x US-09-601-667b-5 ..

Align seg 1/1 to: US-09-601-667b-5 from: 1 to: 256

```
1 GCCAGATTCAATCCCATCNTGTGGAGGCTTCCGCCGCAATTAACAGTGG 50
|||||
169 AlArGpHeAsnPrOllLeuTrPaRgTyrArgGlnAspIleAsnSerGI 185
51 GGAGTCNTCTCCACCAACATGTCATGCTGCAGCTGGAGACGAGTGGG 100
|||||
185 YGlnSerPheLeuProAspMetTyrMetLeuGlnLeuGlnThrSetrPg 202
101 GTCGACATCCACCCAGTCCAGCAGTCCAAAGATGGCATTTTAAATACC 150
|||||
202 YgInGlnSerThrGlnValGlnHisSerThrAspGlyAlaPheAsnAsn 218
151 CAAATTAAGATTGCAGATTTCGCCGGTAATTGTGTGACGNTGACAAATGT 200
|||||
219 ProPheArgLeuAlaIleSerThrGlyAsnPheValThrLeuSerAsnVa 235
201 TCGGACGTGATCTCCAGCTTGGCGATCATGTGTTCGAATGCAGTGTGC 250
|||||
235 lArgSerValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGlnu 252
251 GGCCATTCTCCTCT 264
|||||
252 rGProSerSerSer 256
```

seq\_name: /cgn2\_6/prodata/1/paa/US09\_NEW\_COMB.pep:US-09-601-667b-38

## seq\_documentation\_block:

; Sequence 38, Application US/09601667B  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Peter  
; APPLICANT: Stiefel, Thomas  
; APPLICANT: Voelter, Wolfgang  
; APPLICANT: Welters, Peter  
; TITLE OF INVENTION: Recombinant Mistletoe Lectins  
; FILE REFERENCE: 29841/36636  
; CURRENT APPLICATION NUMBER: US/09/601,667B  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: PCT/EP99/00696  
; PRIOR FILING DATE: 1999-02-03  
; PRIOR APPLICATION NUMBER: D 198 04 210.8  
; NUMBER OF SEQ ID NOS: 41  
; SEQ ID NO 38  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: mistletoe lectin A2  
US-09-601-667b-38

alignment\_scores:  
Quality: 327.00 Length: 88  
Ratio: 4.247 Gaps: 0  
Percent Similarity: 87.500 Percent Identity: 73.864

## alignment\_block:

US-09-627-165b-13 x US-09-601-667b-38 ..

Align seg 1/1 to: US-09-601-667b-38 from: 1 to: 256

```
1 GCCAGATTCAATCCCATCNTGTGGAGGCTTCCGCCGCAATTAACAGTGG 50
|||||
169 AlArGpHeAsnPrOllLeuTrPaRgTyrArgGlnAspIleAsnSerGI 185
51 GGAGTCNTCTCCACCAACATGTCATGCTGCAGCTGGAGACGAGTGGG 100
|||||
185 YGlnSerPheLeuProAspMetTyrMetLeuGlnLeuGlnThrSetrPg 202
101 GTCGACATCCACCCAGTCCAGCAGTCCAAAGATGGCATTTTAAATACC 150
```



TITLE OF INVENTION: Recombinant Mistletoe Lectins  
FILE REFERENCE: 29841/3636  
CURRENT APPLICATION NUMBER: US/09/601,667B  
CURRENT FILING DATE: 2000-10-06  
PRIORITY APPLICATION NUMBER: PCT/EP99/00696  
PRIORITY FILING DATE: 1999-02-03  
PRIORITY APPLICATION NUMBER: D 198 04 210.8  
PRIORITY FILING DATE: 1998-02-03  
NUMBER OF SEQ ID NOS: 41  
SEQ ID NO 2  
LENGTH: 255  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE: Artificial Sequence  
OTHER INFORMATION: M1A-chain  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 15  
OTHER INFORMATION: product= "Xaa is Asp or Glu"  
OTHER INFORMATION: /label= Xaa1  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 63  
OTHER INFORMATION: product= "Xaa is Gly or Gln"  
OTHER INFORMATION: /label= Xaa2  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 66  
OTHER INFORMATION: product= "Xaa is Ile or Val"  
OTHER INFORMATION: /label= Xaa3  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 75  
OTHER INFORMATION: product= "Xaa is Leu or Ala"  
OTHER INFORMATION: /label= Xaa4  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 107  
OTHER INFORMATION: product= "Xaa is missing"  
OTHER INFORMATION: /label= Xaa5  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 113  
OTHER INFORMATION: product= "Xaa is Asn or Thr"  
OTHER INFORMATION: /label= Xaa6  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 117  
OTHER INFORMATION: product= "Xaa is Pro or Thr"  
OTHER INFORMATION: /label= Xaa7  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 134  
OTHER INFORMATION: product= "Xaa is Asp or Glu"  
OTHER INFORMATION: /label= Xaa8  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 141  
OTHER INFORMATION: product= "Xaa is Ser or Thr"  
OTHER INFORMATION: /label= Xaa9  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 145  
OTHER INFORMATION: product= "Xaa is Phe or Tyr"  
OTHER INFORMATION: /label= Xaa10  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 152  
OTHER INFORMATION: product= "Xaa is Thr or Ala"  
OTHER INFORMATION: /label= Xaa11  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 177

OTHER INFORMATION: product= "Xaa is Ala or Tyr"  
OTHER INFORMATION: /label= Xaa12  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 180  
OTHER INFORMATION: product= "Xaa is Tyr or Asp"  
OTHER INFORMATION: /label= Xaa13  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 185  
OTHER INFORMATION: product= "Xaa is Ala or Glu"  
OTHER INFORMATION: /label= Xaa14  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 191  
OTHER INFORMATION: product= "Xaa is Val or Met"  
OTHER INFORMATION: /label= Xaa15  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 219  
OTHER INFORMATION: product= "Xaa is Ile or Phe"  
OTHER INFORMATION: /label= Xaa16  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 224  
OTHER INFORMATION: product= "Xaa is Pro or Ser"  
OTHER INFORMATION: /label= Xaa17  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 225  
OTHER INFORMATION: product= "Xaa is Pro or Thr"  
OTHER INFORMATION: /label= Xaa18  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 232  
OTHER INFORMATION: product= "Xaa is Thr or Ser"  
OTHER INFORMATION: /label= Xaa19  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 236  
OTHER INFORMATION: product= "Xaa is Asp or Ser"  
OTHER INFORMATION: /label= Xaa20  
US-09-601-667B-2

alignment scores:                      Length:                      88  
Quality:                      304.00  
Ratio:                      4.222  
Percent Similarity:                      81.818                      Percent Identity:                      69.318

alignment\_block:  
US-09-627-165B-13 x US-09-601-667B-2                      ..  
Align seg 1/1 to: US-09-601-667B-2 from: 1 to: 255

1 GCCAGATTCATCCATCCATGTCGAGGCTTCCGCCGCAATTACAGTGG 50  
|||||  
168 AlarqphesnpProileutrparg\*\*ArgGln\*\*IleasnSergcl 184  
51 GGAAGTCATTCACCAACATGTCGAGTGCAGTGCAGAGTGG 100  
|||||  
184 y\*\*SerPheleuproasp\*\*TyrtleuGluuGluThrsertRPG 201  
101 GTGCAATCCACCAAGTCCAGTCCAGAGTGCAGATTTTAAATAC 150  
|||||  
201 TgInginserThgInvaIgInIsertHraspelyAlPheasn 217  
151 CAAATTAAGATTGCAGATTCCGCGGTAATTTGTCAGATGCAATGT 200  
|||||  
218 Pro\*\*ArgleuAlaIle\*\*\*GlyAsnPhelValThleu\*\*AsnVa 234  
201 TCCGAGCTGATTCAGAGTTCGAGTGCAGATCTTGTTCGAATGCAATGTC 250  
|||||

LOCATION: (178)..(178

seq\_name: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:US-10-

```
seq_name: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:US-10-083-336A-1
seq_documentation_block:
```



Sequence 1, Application US/10083336A  
GENERAL INFORMATION:  
APPLICANT: Olson, Mark A  
APPLICANT: Miliard, Charles B  
APPLICANT: Byrne, Michael P  
APPLICANT: Wannemacher, Robert W  
TITLE OF INVENTION: R1cin Vaccine and Methods of Making and Using Thereof  
FILE REFERENCE: P67452050 (RIID 01-58)  
CURRENT APPLICATION NUMBER: US/10/083,336A  
CURRENT FILING DATE: 2002-05-21  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 576  
TYPE: PRT  
ORGANISM: R1cinus communis  
US-10-083-336A-1

alignment\_scores:                      Length:    119  
Quality:    178.00                      Gaps:    1  
Ratio:    2.342  
Percent Similarity:    63.866            Percent Identity:    36.975

alignment\_block:  
US-09-627-165b-13 x US-10-083-336A-1 ..

Align seq 1/1 to: US-10-083-336A-1 from: 1 to: 576

```
1 GCCAGATTCAATCCCATCTGTCGAGGCTTGGCCGCAATTAACAGTCG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214 Alarphheglntyrilleglmgumetargrtharqileatgtyras 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 GGAGTCCTCCACCAACATGTACATGCTGAGCTGGAGAGAGTGGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
210 nargarsteralaprosaproservaliethrleugluansetrtpg 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 GTCGACATCCACCCAGTCACAGCTCCAGAGATGGCATTTTAAATACC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
247 lyarglsesthrAlaleglngluserasnglnclyalaphelaser 263
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 CAATAAGATTGCAGATTTCGCGGGAATCTTGACGNTGACCAATGT 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 ProtleaglntleuglnargargasnglyserlysphesterValtyraspya 280
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 TCGCAGCTGATCTCCAGCTTGGGATCATGTTGCGAATGACATGCTC 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
280 lserlleuleuleprolleilealaleumetvaltyrargysalaprop 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 GGCATTCCTCTCTCTGACACACACCTTGGCGGCTGCTCAAGTCCGTC 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 robrsersterglnpheser.....leuleulleargproval 309
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 GTGATGCGCGCCACAGATGTCACCTGCACTTTTCCGACACCCACCTGCG 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
310 ValpProasnpheasnalaspyvalCysmetaspproglnuProillevalr 326
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
351 CATCGTA 357.
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
326 gileval 328
```

seq\_name: /cgn2\_6/ptodata/1/pae/us10\_NEW\_COMB pep: US-10-127-890-1

seq\_documentation\_block:  
Sequence 1, Application US/10127890

GENERAL INFORMATION:

APPLICANT: Belter, Marc D  
Carroll, Stephen F.

Studinka, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

Proteins

NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-MAY-1996

APPLICATION NUMBER: ECT/US94/05348

FILING DATE: 12-MAY-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-9155

TELEFAX: 312/707-9158

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-127-890-1.

alignment\_scores:

Quality:    138.00                      Length:    87  
Ratio:    2.379                      Gaps:    0  
Percent Similarity:    66.667            Percent Identity:    35.632

alignment\_block:  
US-09-627-165b-13 x US-10-127-890-1 ..

Align seq 1/1 to: US-10-127-890-1 from: 1 to: 267

```
1 GCCAGATTCAATCCCATCTGTCGAGGCTTGGCCGCAATTAACAGTCG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179 Alarphheglntyrilleglmgumetargrtharqileatgtyras 195
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
195 nargarsteralaprosaproservaliethrleugluansetrtpg 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 GGAGTCCTCCACCAACATGTACATGCTGAGCTGGAGAGAGTGGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
195 nargarsteralaprosaproservaliethrleugluansetrtpg 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 GTCGACATCCACCCAGTCACAGCTCCAGAGATGGCATTTTAAATACC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
212 lyarglsesthrAlaleglngluserasnglnclyalaphelaser 228
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 CAATAAGATTGCAGATTTCGCGGGAATCTTGACGNTGACCAATGT 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
229 ProtleaglntleuglnargargasnglyserlysphesterValtyraspya 245
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 TCGCAGCTGATCTCCAGCTTGGGATCATGTTGCGAATGACATGCTC 250
```

245 |serlleulelprollellealelauleu|valfyrargysala|prop 262  
251 GGCCATTCCTCC 261  
262 roPtoSerSer 265

seq\_name: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:us-10-127-890-7

seq\_documentation\_block:

Sequence 7, Application us/10127890

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

Studulka, Gary M.

Carroll, Stephen F.

Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESSES:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-May-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-May-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-May-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-Dec-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-Jun-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-Nov-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

INFORMATION FOR SEQ ID NO: 7:

LENGTH: 263 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-127-890-7

alignment\_scores:

Quality: 92.00 Length: 82

Ratio: 1.769 Gaps: 3

Percent Similarity: 63.415 Percent Identity: 34.146

alignment\_block:

US-09-627-165b-13 x US-10-127-890-7 ..

Align seg 1/1 to: US-10-127-890-7 from: 1 to: 263

1 GCCAGATTCAATCCCATCTGTGGAGCTTGCCGGCAATTACAGTGC 50  
162 AlaArgPheLysTyrIleGluGlnGlnIleGlnValArgAlaArgAs 178  
51 GGAGTCNTCTCCACCAACATGTACATGCTGAGCTGAGACAGATTGGC 100  
178 pGlu.....ValProSerLeuAlaThrIleSerLeuGlnAsnSerTrp 193  
101 GTGCACATTCACCCCAAGTCCAG.....CAGTCCAAAGATGGCATTTT 144  
193 ergIleuSerLysGlnIleGlnLeuAlaGlnIleGlnAsnGlnYllePhe 209  
145 AATACCCAAATAGATTGACAGATTCCCGCGTACTTGTGACGNTGAG 194  
210 ArgThrProIleValLeuValAspAsnLysGlnLysArgValGlnIleTh 226  
195 CAATGTT.....CGCGACGTGATCTCCAGCTTGGGAGATCATCTTG 234  
226 rAsnValThrSerLysValThrSerAsnIleGlnLeuLeu 241

seq\_name: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:us-10-127-890-9

seq\_documentation\_block:

Sequence 9, Application us/10127890

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

Studulka, Gary M.

Carroll, Stephen F.

Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESSES:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-May-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-May-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-May-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-Dec-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-Jun-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-Nov-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 261 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
US-10-127-890-9

alignment\_scores:  
Quality: 79.00 Length: 59  
Ratio: 2.026 Gaps: 1  
Percent Similarity: 66.102 Percent Identity: 33.698

alignment\_block:

US-09-627-165b-13 x US-10-127-890-9 ..

Align seg 1/1 to: US-10-127-890-9 from: 1 to: 261

1 GCCAGATTCAATCCCATGTCGAGCGCTCGCCGCAATTACAGTGG 50  
|||||  
177 AATATGPhelYTYrIleGluAsnGlnValYsThrAsnPhenAsnArgAs 193  
51 GGAGTCNTCTCCACCAACATGTACATGCTGAGCTGAGAGAGAGTGGG 100  
||| |||||  
193 pPheSer.....ProsnAspIysValIeuAspIeuGluAsnTrpG 208  
101 GTGCAATCCACCCAGTCAGCATGTCACAGATGGCATTTTAAATACC 150  
||| |||||  
208 IYsISeThrAlaIleHisAsnSerIysAsnIYAlaIeuProIys 224  
151 CAATTAAGATGTCAGATTTCGCCGCGT 177  
|||||  
225 ProIeuGluIeuIYsAsnAlaAspGly 233

seq\_name: /cgn2\_6/prodata/1/paa/US10\_NEW\_COMB.pep:US-10-127-890-102

seq\_documentation\_block:

Sequence 102, Application US/10127890

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

Carroll, Stephen F.

Studenka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-May-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-May-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-May-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-Dec-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.p4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-9155

TELEFAX: 312/707-8889

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:

LENGTH: 251 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 102:

US-10-127-890-102

alignment\_scores:

Quality: 75.50 Length: 82

Ratio: 1.541 Gaps: 2

Percent Similarity: 59.756 Percent Identity: 28.049

alignment\_block:

US-09-627-165b-13 x US-10-127-890-102 ..

Align seg 1/1 to: US-10-127-890-102 from: 1 to: 251

1 GCCAGATTCAATCCCATGTCGAGCGCTCGCCGCAATTACAGTGG 50  
|||||  
168 AATATGPhelrPhelIleGluAsnGlnIleArgAsn.....AsnPhel 182  
51 GGAGTCNTCTCCACCAACATGTACATGCTGAGCTGAGAGAGAGTGGG 100  
||| |||||  
182 nGlnArgIleArgProAlaAsnAsnThrIleSerIeuGluAsnIYsTrpG 199  
101 GTGCAATCCACCCAGTCAGCATGTCACAGATGGCATTTTAAAT 147  
||| |||||  
199 IYsISeThrAlaIleHisAsnSerIysAsnIYAlaIeuProIys 215  
148 ACCCAATTAAGATGTCAGATTTCGCCGCGTAACTTGTGACGNTGACCA 197  
|||||  
216 GlnAlaValGluIeuGluArgAlaAsnGlyIYsTYrTYrValThrAl 232  
198 TGTTCGCGAGCTGATCTCCAGCTTGCGATCATGTTGTCGAATGC 243  
||| |||||  
232 aValAspGlnValYsProIYsIleAlaIeuIeuIYsPheValYs 247

seq\_name: /cgn2\_6/prodata/1/paa/US10\_NEW\_COMB.pep:US-10-127-890-111

seq\_documentation\_block:

Sequence 111, Application US/10127890

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

Carroll, Stephen F.

Studenka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 368-1248  
INFORMATION FOR SEQ ID NO: 111:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 111:  
US-10-127-890-111

alignment\_scores:  
Quality: 75.50 Length: 82  
Ratio: 1.541 Gaps: 2  
Percent Similarity: 59.756 Percent Identity: 28.049

## alignment\_block:

US-09-627-165b-13 x US-10-127-890-111 ..

Align seg 1/1 to: US-10-127-890-111 from: 1 to: 251

```
1 GCCGATTCAATCCCATCCTGTGAGGCTTCGCCGCCCAATTAAACAGTGG 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
168 ALaArgpHeThrPheIleGluAsnGlnIleArgAsn.....AspPheG1 182
51 GGAGTCNTCTCCACCAACATGTACATGCTGCGAGTGGAGAGCATGGGG 100
   :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 nGlnArgIleArgProAlaAsnAsnThrIleSerLeuGluAsnLysTrpG 199
101 GTGCGCAATCCCAACCAAGTCACAGATGC...AAGATGGCATTTTAA 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
199 LysIleuSerPheGlnIleArgThrSerGlyAlaAsnIleuMetPheSer 215
148 ACCCAATATAGATTCAGATTCGCCGCCGCTTACTTGTGACGNTGACGAA 197
   :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
216 G1AlaIaValGluLeuGluArgAlaAsnGlyLysTyrTyrValThrAl 232
198 TGTTCGCGAGTGCATTCACCTTGCGGATGCATGATGTGTTCGATATGC 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
232 aValAspGlnValLysProLysIleAlaLeuLeuLysPheValCys 247
```

OM of: US-09-627-165b-15 to: Pending\_Patents\_AA\_Main:\* out\_format: pfs  
Date: Jul 30, 2002 4:03 PM  
About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

```
-MODEL=franklin.n2p.model -DEV=xlp  
-O=/cgn2_6/ptodata/2/paa/US096/COMB.pep:US-09-627-165b-2 + 620.00 1236.53 1.3e-60  
-DB=Pending_Patents_AA_Main -FASTA=fastac -SUFFIX=rapm  
-GAPOP=12.000 -GAPOP=4.000 -MISMATCH=0.100 -LOOPCL=0.000  
-LOOEXT=0.000 -GAPOP=4.500 -MISMATCH=0.050 -GAPOP=10.000  
-GAEXT=0.500 -GAPOP=6.000 -GAEXT=7.000 -GAPOP=10.000  
-GAEXT=0.500 -DELOP=6.000 -DEEXT=7.000 -START=1  
-MATRIX=blomsm62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200  
-THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09627165 @CGNL_1.634 -NCPU=6 -LONCLOG  
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPHY -WAIT -THREADS=1
```

## Search information block:

```
Query: US-09-627-165b-15  
Query Length: 522  
Database: Pending_Patents_AA_Main:*  
Database sequences: 3502263  
Database length: 351980561  
Search time (sec): 417.700000
```

## score\_list:

```
Sequence      Strd Orig      ZScore      EScore      Len      Documentation  
/cgn2_6/ptodata/2/paa/US096/COMB.pep:US-09-627-165b-2 + 620.00 1236.53 1.3e-60 254  
/cgn2_6/ptodata/2/paa/US093/COMB.pep:US-09-347-064-8 + 619.00 1234.59 1.6e-60 252  
/cgn2_6/ptodata/2/paa/US093/COMB.pep:US-09-347-064-2 + 615.00 1226.49 4.6e-60 252  
/cgn2_6/ptodata/2/paa/US093/COMB.pep:US-09-627-165b-4 + 615.00 1226.49 4.6e-60 252  
/cgn2_6/ptodata/2/paa/US096/COMB.pep:US-09-627-165b-14 + 611.00 1218.22 1.3e-59 256  
/cgn2_6/ptodata/2/paa/US096/COMB.pep:US-09-627-165b-1 + 408.50 816.02 7.1e-37 119  
/cgn2_6/ptodata/2/paa/US095/COMB.pep:US-09-538-873-3 + 243.00 472.95 4.3e-18 251  
/cgn2_6/ptodata/2/paa/US095/COMB.pep:US-09-668-419-3 + 243.00 472.95 4.3e-18 251  
/cgn2_6/ptodata/2/paa/US07/COMB.pep:US-07-787-567A-1 + 217.00 419.64 3.8e-15 267  
/cgn2_6/ptodata/2/paa/US084/COMB.pep:US-08-406-801-34 + 217.00 419.64 3.8e-15 267  
/cgn2_6/ptodata/2/paa/US095/COMB.pep:US-09-538-873-1 + 217.00 419.64 3.8e-15 267  
/cgn2_6/ptodata/2/paa/US096/COMB.pep:US-09-610-838-1 + 217.00 419.64 3.8e-15 267  
/cgn2_6/ptodata/2/paa/US096/COMB.pep:US-09-668-419-1 + 217.00 419.64 3.8e-15 267  
/cgn2_6/ptodata/2/paa/US097/COMB.pep:US-08-378-761-27 + 213.00 410.67 1.1e-14 289  
/cgn2_6/ptodata/2/paa/US083/COMB.pep:US-08-263-261-5 + 167.50 318.53 1.5e-09 289  
/cgn2_6/ptodata/2/paa/US082/COMB.pep:US-08-263-261-6 + 158.50 301.94 1.5e-08 247  
/cgn2_6/ptodata/2/paa/US07/COMB.pep:US-07-504-775-9 + 156.50 296.25 2.6e-08 289  
/cgn2_6/ptodata/2/paa/US07/COMB.pep:US-07-739-143-4 + 156.50 296.25 2.6e-08 289  
/cgn2_6/ptodata/2/paa/US082/COMB.pep:US-08-263-261-9 + 156.50 296.25 2.6e-08 289  
/cgn2_6/ptodata/2/paa/US084/COMB.pep:US-08-480-432-4 + 156.50 296.25 2.6e-08 289  
/cgn2_6/ptodata/2/paa/US091/COMB.pep:US-09-120-523-19 + 155.50 295.87 3.2e-08 247  
/cgn2_6/ptodata/2/paa/US093/COMB.pep:US-09-360-242-39 + 155.50 295.87 3.2e-08 247  
/cgn2_6/ptodata/2/paa/US087/COMB.pep:US-08-715-074A-2 + 155.50 295.87 3.2e-08 247  
/cgn2_6/ptodata/2/paa/US087/COMB.pep:US-07-787-567A-7 + 154.50 293.18 4.2e-08 263  
/cgn2_6/ptodata/2/paa/US096/COMB.pep:US-09-610-838-7 + 154.50 293.18 4.2e-08 263  
/cgn2_6/ptodata/2/paa/US097/COMB.pep:US-09-711-485-6 + 153.50 291.81 5.4e-08 247  
/cgn2_6/ptodata/2/paa/US096/COMB.pep:US-09-610-838-6 + 153.50 291.81 5.4e-08 247  
/cgn2_6/ptodata/2/paa/US097/COMB.pep:US-09-711-485-6 + 153.50 291.81 5.4e-08 247  
/cgn2_6/ptodata/2/paa/US07/COMB.pep:US-07-787-567A-6 + 152.50 289.21 7.1e-08 261  
/cgn2_6/ptodata/2/paa/US096/COMB.pep:US-09-610-838-9 + 152.50 289.21 7.1e-08 261  
/cgn2_6/ptodata/2/paa/US097/COMB.pep:US-09-711-485-9 + 152.50 289.21 7.1e-08 261
```

```
/cgn2_6/ptodata/2/paa/US07/COMB.pep:US-07-504-775-4 + 151.00 286.71 1.0e-07 248  
/cgn2_6/ptodata/2/paa/US082/COMB.pep:US-08-263-261-4 + 151.00 286.71 1.0e-07 248  
/cgn2_6/ptodata/2/paa/US07/COMB.pep:US-07-504-775-3 + 151.00 285.74 1.1e-07 272  
/cgn2_6/ptodata/2/paa/US082/COMB.pep:US-08-263-261-3 + 151.00 285.74 1.1e-07 272
```

## seq\_documentation\_block:

```
Sequence 2 Application US/096271558  
GENERAL INFORMATION:  
APPLICANT: KIM, Jong-Bae  
TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM album coloratum, AND PROTEINS  
FILE REFERENCE: AND LECTINS ISOLATED THEREFROM  
CURRENT FILING DATE: 2000-07-27  
NUMBER OF SEQ ID NOS: 16  
SEQ ID NO 2  
LENGTH: 254  
TYPE: PRT  
ORGANISM: Viscum album coloratum  
FEATURE:  
NAME/KEY: misc.feature  
US-09-627-165b-2
```

## alignment\_scores:

```
Quality: 620.00 Length: 140  
Ratio: 4.593 Gaps: 0  
Percent Similarity: 96.429 Percent Identity: 87.857
```

## alignment\_block:

```
US-09-627-165b-15 x US-09-627-165b-2  
Align seq 1/1 to: US-09-627-165b-2 from: 1 to: 254
```

```
1 TATACGATCTGGACCGATTCGATCCGATCAGGACGACGATCCCTGCGG 50  
115 TTTTTHAATGAGTCTGATCCTGATCCTGATCCTGATCCTGATCCTG 131  
51 TATACGATCTGGACCGATTCGATCCGATCAGGACGACGATCCCTGCGG 100  
1 TATACGATCTGGACCGATTCGATCCGATCAGGACGACGATCCCTGCGG 100  
131 TATACGATCTGGACCGATTCGATCCGATCAGGACGACGATCCCTGCGG 148  
101 GCACCGGCGGACGATTCGATCCGATCAGGACGACGATCCCTGCGG 150  
148 TTTTTHAATGAGTCTGATCCTGATCCTGATCCTGATCCTGATCCTG 164  
151 GATCGGCGGACGATTCGATCCGATCAGGACGACGATCCCTGCGG 200  
165 GATCGGCGGACGATTCGATCCGATCAGGACGACGATCCCTGCGG 181  
201 GATCGGCGGACGATTCGATCCGATCAGGACGACGATCCCTGCGG 250  
181 TATACGATCTGGACCGATTCGATCCGATCAGGACGACGATCCCTGCGG 198  
251 GATCGGCGGACGATTCGATCCGATCAGGACGACGATCCCTGCGG 300  
198 TTTTTHAATGAGTCTGATCCTGATCCTGATCCTGATCCTGATCCTG 214  
301 TATACGATCTGGACCGATTCGATCCGATCAGGACGACGATCCCTGCGG 350  
215 TATACGATCTGGACCGATTCGATCCGATCAGGACGACGATCCCTGCGG 231  
351 TATACGATCTGGACCGATTCGATCCGATCAGGACGACGATCCCTGCGG 400  
231 TATACGATCTGGACCGATTCGATCCGATCAGGACGACGATCCCTGCGG 248  
401 GATCGGCGGACGATTCGATCCGATCAGGACGACGATCCCTGCGG 420  
248 TATACGATCTGGACCGATTCGATCCGATCAGGACGACGATCCCTGCGG 254  
seq_name: /cgn2_6/ptodata/2/paa/US093/COMB.pep:US-09-347-064-8
```

```
seq_documentation_block:
; Sequence 8, Application US/09347064
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE OF INVENTION: album
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
; US-09-347-064-8
```

```
alignment_scores:
Quality: 619.00 Length: 138
Ratio: 4.654 Gaps: 0
Percent Similarity: 96.377 Percent Identity: 86.232
```

alignment\_block:

US-09-627-165b-15 x US-09-347-064-8 ..

Align seg 1/1 to: US-09-347-064-8 from: 1 to: 252

```
1 TACACGATCTGAGCGATACCGCGTCATAGGACACCATCCCTCTGG 50
|||||
115 TyrProaspLeuGlnArgTyrAlaGlyHisArgAspGlnIleProLeuG1 131
51 TATAGAGAGTCATTCATCCGTCCTCGCGCTTCGTTATCCAGCGGCA 100
|||||
131 YtleAspGlnIleuIleGlnSerValThrAlaLeuArgPheProGlyG1s 148
101 GCACCGCGGCCAAGCTCGTCCCTTATATCCTCATTCAGATGATCTCC 150
|||||
148 erThrArgThrGlnAlaArgSerIleLeuIleGlnMetIleSer 164
151 GAGCGCGGAGATTCATCCGATCTTTGGAGGCGTCGCAATACATTA 200
|||||
165 GluAlaAlaArgPheAsnProIleLeuTrpArgAlaArgGlnTyrIleAs 181
201 CAGCGGGAGTCATTTCTCCGACATGTCATGCTCCAGCTGGAGACTA 250
|||||
181 nSerGlyAlaSerPheLeuProAspValTyrMetLeuGlnLeuGluThrS 198
251 GTTGGGGCAACATCCACGCAAGTCCAGAGCTTACGATCGCGTTT 300
|||||
198 erTrpGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 214
301 AATAACCATTTGGGTGGTATATCCACCGGTAATCTGTAGACTTGAG 350
|||||
215 AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValThrLeuTh 231
351 CAATGTCGCGAGTCATCGGACCTTAGGATCATGCTGTTGAGACTA 400
|||||
231 rAsnValAlaArgAspValIleAlaSerLeuAlaIleMetLeuPheValCySg 248
401 GGGACCGACCATCT 414
|||||
248 YgluArgProSer 252
```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-347-064-8

```
seq_documentation_block:
; Sequence 8, Application US/09347064A
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE OF INVENTION: album
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
; US-09-347-064-8
```

```
alignment_scores:
Quality: 619.00 Length: 138
Ratio: 4.654 Gaps: 0
Percent Similarity: 96.377 Percent Identity: 86.232
```

alignment\_block:

US-09-627-165b-15 x US-09-347-064-8 ..

Align seg 1/1 to: US-09-347-064-8 from: 1 to: 252

```
1 TACACGATCTGAGCGATACCGCGTCATAGGACACCATCCCTCTGG 50
|||||
115 TyrProaspLeuGlnArgTyrAlaGlyHisArgAspGlnIleProLeuG1 131
51 TATAGAGAGTCATTCATCCGTCCTCGCGCTTCGTTATCCAGCGGCA 100
|||||
131 YtleAspGlnIleuIleGlnSerValThrAlaLeuArgPheProGlyG1s 148
101 GCACCGCGGCCAAGCTCGTCCCTTATATCCTCATTCAGATGATCTCC 150
|||||
148 erThrArgThrGlnAlaArgSerIleLeuIleGlnMetIleSer 164
151 GAGCGCGGAGATTCATCCGATCTTTGGAGGCGTCGCAATACATTA 200
|||||
165 GluAlaAlaArgPheAsnProIleLeuTrpArgAlaArgGlnTyrIleAs 181
201 CAGCGGGAGTCATTTCTCCGACATGTCATGCTCCAGCTGGAGACTA 250
|||||
181 nSerGlyAlaSerPheLeuProAspValTyrMetLeuGlnLeuGluThrS 198
251 GTTGGGGCAACATCCACGCAAGTCCAGAGCTTACGATCGCGTTT 300
|||||
198 erTrpGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 214
301 AATAACCATTTGGGTGGTATATCCACCGGTAATCTGTAGACTTGAG 350
|||||
215 AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValThrLeuTh 231
351 CAATGTCGCGAGTCATCGGACCTTAGGATCATGCTGTTGAGACTA 400
|||||
231 rAsnValAlaArgAspValIleAlaSerLeuAlaIleMetLeuPheValCySg 248
401 GGGACCGACCATCT 414
|||||
248 YgluArgProSer 252
```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-347-064-2

```
seq_documentation_block:
; Sequence 2, Application US/09347064
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; TITLE OF INVENTION: album
; FILE REFERENCE: 09282-5
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
; US-09-347-064-2
```

```
alignment_scores:
Quality: 615.00 Length: 137
Ratio: 4.659 Gaps: 0
Percent Similarity: 96.350 Percent Identity: 86.131
```

alignment\_block:
US-09-627-165b-15 x US-09-347-064-2 ..

Align seg 1/1 to: US-09-347-064-2 from: 1 to: 252

```
1 TACACAGATCTGAGCGATACCGCGGTATAGGACGATCCCTCTGGG 50
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116 TTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 132
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 TATAGAGAACTCATTCATCCGCTCTGCGGCTTATTCAGCGCGCA 100
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 YTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 149
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 GCACCGCGGCGCAAGCTCTCTCTTATTCATTCAGATGATGATGAT 150
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 erThrAlgrhGlnAlaArgSerIleuIleuIleuIleuIleuIleu 165
151 GAGCGCGGAGATTCATTCATCCGCTCTGCGGCTTATTCAGCGCA 200
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
166 GlnAlaAlaArgPheAsnProIleuTrpArgAlaArgGlnTyrIleAs 182
201 CACCGCGGAGATTCATTCATCCGCTCTGCGGCTTATTCAGCGCA 250
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 nserGlyAlaSerPheIleuProAspValTyrMetLeuGlnLeuGln 199
251 GTTGGGCGCAACAATCCAGCAAGTCCAGAGTTCAGCGATGCGGTT 300
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 erTrpGlyGlnGlnSerThrGlnValGlnIleAsSerThrAspGlyAlaPhe 215
301 AATAACCATTTGGTGGGTATATCCACCGGTATACCTTCGAGCGT 350
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 AsnAsnProIleuArgLeuAlaIleProGlyAsnPheValThrLeuTh 232
351 CAATGTCGCGAGTGTATCCGCAAGTTCAGCGATGATGATGATGAT 400
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232 rAsnValArgAspValIleIleAlaSerLeuAlaIleMetLeuPheValCysG 249
401 GGGACCGACCA 411
:::|||||
249 LysGlnArgPro 252
```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-347-064-2

```
seq_documentation_block:
; Sequence 2, Application US/09347064A
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; TITLE OF INVENTION: album
; FILE REFERENCE: 09282-5
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
; US-09-347-064-2
```

```
alignment_scores:
Quality: 615.00 Length: 137
Ratio: 4.659 Gaps: 0
Percent Similarity: 96.350 Percent Identity: 86.131
```

alignment\_block:
US-09-627-165b-15 x US-09-347-064-2 ..

Align seg 1/1 to: US-09-347-064-2 from: 1 to: 252

```
1 TACACAGATCTGAGCGATACCGCGGTATAGGACGATCCCTCTGGG 50
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116 TTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 132
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 TATAGAGAACTCATTCATCCGCTCTGCGGCTTATTCAGCGCGCA 100
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 YTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 149
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 GCACCGCGGCGCAAGCTCTCTCTTATTCATTCAGATGATGATGAT 150
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 erThrAlgrhGlnAlaArgSerIleuIleuIleuIleuIleuIleu 165
151 GAGCGCGGAGATTCATTCATCCGCTCTGCGGCTTATTCAGCGCA 200
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
166 GlnAlaAlaArgPheAsnProIleuTrpArgAlaArgGlnTyrIleAs 182
201 CACCGCGGAGATTCATTCATCCGCTCTGCGGCTTATTCAGCGCA 250
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 nserGlyAlaSerPheIleuProAspValTyrMetLeuGlnLeuGln 199
251 GTTGGGCGCAACAATCCAGCAAGTCCAGAGTTCAGCGATGCGGTT 300
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 erTrpGlyGlnGlnSerThrGlnValGlnIleAsSerThrAspGlyAlaPhe 215
301 AATAACCATTTGGTGGGTATATCCACCGGTATACCTTCGAGCGT 350
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 AsnAsnProIleuArgLeuAlaIleProGlyAsnPheValThrLeuTh 232
351 CAATGTCGCGAGTGTATCCGCAAGTTCAGCGATGATGATGATGAT 400
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232 rAsnValArgAspValIleIleAlaSerLeuAlaIleMetLeuPheValCysG 249
401 GGGACCGACCA 411
:::|||||
249 LysGlnArgPro 252
```

seq\_name: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:US-09-627-165b-4

## seq\_documentation\_block:

```
; Sequence 4, Application US/09627165B
; GENERAL INFORMATION:
; APPLICANT: Kim, Jong-Bae
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM album coloratum, AND PROTEINS
; FILE REFERENCE: AND LECTINS ISOLATED THEREFROM
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 4
; LENGTH: 254
; TYPE: PRP
; ORGANISM: Viscum album coloratum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 240
; OTHER INFORMATION: Xaa = any amino acid
US-09-627-165B-4
```

## alignment\_scores:

```
Quality: 614.00 Length: 140
Ratio: 4.582 Gaps: 0
Percent Similarity: 95.714 Percent Identity: 86.429
```

## alignment\_block:

```
US-09-627-165B-15 x US-09-627-165B-4 ..
```

```
Align seg 1/1 to: US-09-627-165B-4 from: 1 to: 254
```

```
1 TACACGATCTGAGCGATACGCCGGTCATAGAGACCAATCCCTCTGG 50
115 TYRPROASPLEUGLARGTYRALAGLYHISARGSPGHLLEPROLEGL 131
51 TATPAGAGACTCATTCATCCGTCGCGCTGCTATCCAGAGCGGCA 100
131 YLLESPGHLNEULLEGLSERVALSERLALENATGPHEPROGLYSER 148
101 GCACCGCGGCCCAAGCTCGTCCCTVMAATCCATTCAGATGATCTCC 150
148 snthArAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 164
151 GAGCGCGGAGATTCATCCATCTTTTGGAGGCGTCGCCAATACATTA 200
165 GIUALAAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 181
201 CAGCGGGAGATCATTCCTCCCGACATGACATGCTCGACGAGACTA 250
181 rserllyglserPheleuPProAspHrTYrIleuGlnleuGlnuHrS 198
251 GTTGGGCGCAACATCCACGCAAGTCACAGACTTCACGAGATGGCTT 300
198 eTrTgElYglnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 214
301 AATAACCATTTGGTGGTATATCCACCGGTAACCTTCGACGCTGAG 350
215 AsnaAnProlleArGleuThrIleSerThnGlyValPheValThrLeu 231
351 CAATGTGCGAGCGATCGCCAGCTTAGATCATCTGCTGTATGAGTA 400
231 rAsnValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 248
401 GGGACCGACCATCTCTCC 420
248 lAspArgPProSerSerSer 254
```

```
seq_name: /cgn2_6/ptodata/2/paa/US096_COMB.pep:US-09-627-165B-6
```

## seq\_documentation\_block:

```
; Sequence 6, Application US/09627165B
; GENERAL INFORMATION:
```

## APPLICANT: Kim, Jong-Bae

```
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM album coloratum, AND PROTEINS
; FILE REFERENCE: AND LECTINS ISOLATED THEREFROM
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRP
; ORGANISM: Viscum album coloratum
; FEATURE:
; NAME/KEY: misc_feature
US-09-627-165B-6
```

## alignment\_scores:

```
Quality: 611.00 Length: 139
Ratio: 4.629 Gaps: 0
Percent Similarity: 94.964 Percent Identity: 86.331
```

## alignment\_block:

```
US-09-627-165B-15 x US-09-627-165B-6 ..
```

```
Align seg 1/1 to: US-09-627-165B-6 from: 1 to: 256
```

```
1 TACACGATCTGAGCGATACGCCGGTCATAGAGACCAATCCCTCTGG 50
117 TYRALAAsPLeuGlnLARGTYRALAGLYHISARGSPARGILEPROLEUG 133
51 TATPAGAGACTCATTCATCCGTCGCGCTGCTATCCAGAGCGGCA 100
133 YARGLUPROLEULLEARGSERVALSERLALENATGPYRPROGLYIS 150
101 GAGCGCGGCCCAAGCTCGTCCCTATTAACCCATCATACATATCC 150
150 eTrThArAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 166
151 GAGCGCGGAGATTCATCCATCTTTTGGAGGCGTCGCCAATACATTA 200
167 GIUALAAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 183
201 CAGCGGGAGATCATTCCTCCGACATGATACGCTCGAGCTGAGACTA 250
183 nserlYValserTYrleuPProAspValnYrmetleuGlnleuGlnuH 200
251 GTTGGGCGCAACATCCACGCAAGTCACAGACTTCACGAGATGGCTT 300
200 eTrTgElYglnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 216
301 AATAACCATTTGGTGGTATATCCACCGGTAACCTTCGACGCTGAG 350
217 AsnaAnProlleArGleuThrIleSerThnGlyValPheValThrLeu 233
351 CAATGTGCGAGCGATCGCCAGCTTAGATCATCTGCTGTATGAGTA 400
233 rAsnValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 250
401 GGGACCGACCATCTCTCC 417
250 rAspArgPProSerSerSer 255
```

```
seq_name: /cgn2_6/ptodata/2/paa/US096_COMB.pep:US-09-627-165B-14
```

## seq\_documentation\_block:

```
; Sequence 14, Application US/09627165B
; GENERAL INFORMATION:
; APPLICANT: Kim, Jong-Bae
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM album coloratum, AND PROTEINS
; FILE REFERENCE: AND LECTINS ISOLATED THEREFROM
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
```



Wed Jul 31 08:46:43 2002

us-09-627-165b-15.ram

Page 5

```
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 14
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Viscum album coloratum
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 7, 64, 111
; OTHER INFORMATION: xaa - any amino acid
US-09-627-165B-14

alignment_scores:
  Quality: 408.50      Length: 122
  Gaps: 1
  Ratio: 4.045
  Percent Similarity: 82.787      Percent Identity: 69.672

alignment_block:
US-09-627-165B-15 x US-09-627-165B-14 ..
Align seg 1/1 to: US-09-627-165B-14 from: 1 to: 119

157 GCGAGATTCATCCATCTTTGGAGGGCTGCGCAATACATTAACACGG 206
1 AlatrphensnProle**TtrpargleuArGARglnleasnSerGl 17
207 GGAGTCATTTCTCCGACATGTACATGCTCGAGCTGGAGACTGTTGGG 256
17 ygluSerProProsnmetyrMetleuGlnleuGlnthrsertpg 34
257 GCCAATATCCAGCAGTCCAGCAGTTCACGATGCGCTTTTAAATAC 306
34 lYArGlnSerThnGlnValGlnGlnSerlysnpelyllephasnThr 50
307 CCATTCGGTGGGTATATCCACCGGTAACTGTGACGCTGAGCAATGT 356
51 GlnleArGlnGlnleSerleAGlYasnphValThr**SerasnVa 67
357 TCCCGACGCGATCCGACGCTTAGCATCATGTTGTTGTTAGAGGACC 406
67 lArGAspVallleSerSerleuAlalleMetleuPhelGluCysSerGly 84
407 GACCATCTCTCCGACGCTGCGCTATTGGCGCGTGCATACACCCGTC 456
84 rGrProPheserSerleuAspHisProserProleuLeuLeuArGSerVal 100
457 TTGGAATAATAGCGCGCGCTGACGATGTTACTGCTGACTGCTCCGACC 506
101 ValAsp.....AlaAlaAsnAspValThrCysThr**SerGluPr 114
507 CACCGTGGCGATCGTA 522
114 othrValArGlyleVal 119

seq_name: /cgn2_6/ptodata/2/paa/US095_COMB.pep:US-09-538-873-3

seq_documentation_block:
; Sequence 3, Application US/09538873
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
; TITLE OF INVENTION: SYNDROME (VLS)
; FILE REFERENCE: US/09/538,873
; CURRENT APPLICATION NUMBER: US/09/538,873
; CURRENT FILING DATE: 2000-03-30
; EARLIER APPLICATION NUMBER: 60/126,826
; EARLIER FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
```

```
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Abrus precatorius
; US-09-538-873-3

alignment_scores:
  Quality: 243.00      Length: 145
  Gaps: 3
  Ratio: 2.505
  Percent Similarity: 66.897      Percent Identity: 40.000

alignment_block:
US-09-627-165B-15 x US-09-538-873-3 ..
Align seg 1/1 to: US-09-538-873-3 from: 1 to: 251

1 TACACAGATCTGGAGCATAGCGCGGTACAT...AGGACAGATCCCTCT 47
111 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 TGTGlyAspLeuGlnArGlyrPrAlaHnGlnSerArGlnGlnlleProle 129
48 GGTATAGAGAACTCATTCATCCGCTCTGCGCGCTTCGTTATCCAGCG 97
129 uGlyLeuGlnAlaLeuThnHnGlylleSerPheArGSerGlyGly 146
98 GCACACCCGCGCCAGCTGCTTCCCTATATATCTCATTCAGATATC 147
146 snAspAsnGlnGlnYsAlaArGlyrleuAllelleGlnMetVal 162
148 TCCGAGCGCGGATTCATATCCCATCTTTGGAGGCGCTGCCAATACAT 197
163 AlaAlaAlaAlaArGlyrPheArGlyrleSerAsnArGValArGValSerIl 179
198 TAACAGCGGGAGATTCCTCCGACATGTACATGCTGACGTGAGGA 247
179 eGlnThrleYThrAlaPheGlnProAspAlaAlaMetleSerleuGln 196
248 CTAATGGGGGCCAACAATCCACCAAGCTCAGCAGATGAGGATGCGCTT 297
196 snAsnTrpAspAsnleuSerArGlyValGlnleuSerValGlnsPhr 212
298 TTTATATACCATTTGCTGGGTATATCCACCGGTAACTGCTGACCTT 347
213 PheProAsnGln.....ValThrle 219
348 GACCATGTTTCGAC.....G 364
219 uThrAsnleArGAsnGlnProVallleValAspSerleuSerHisProT 236
365 TGATCCCGACGCTTAGCATCATGTTGTTGATGT 399
236 hrValAlaValleuAlaLeuMetleuPheValCys 247

seq_name: /cgn2_6/ptodata/2/paa/US096_COMB.pep:US-09-668-419-3

seq_documentation_block:
; Sequence 3, Application US/09668419
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN E.
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: RIGIN A CHAIN MUTANTS LACKING ENZYMATIC ACTIVITY AS VACCINES
; TITLE OF INVENTION: AGAINST AEROSOLIZED RIGIN
; FILE REFERENCE: US/09/668,419
; CURRENT APPLICATION NUMBER: US/09/668,419
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Abrus precatorius
; US-09-668-419-3
```

```

alignment_scores:
  Quality: 243.00      length: 145
  Ratio: 2.505         Gaps: 3
Percent Similarity: 66.697      Percent Identity: 40.00000
alignment block:
US-09-627-165B-15 x US-09-666-419-3  ..

```

Align seg 1/1 to: US-09-668-419-3 from: 1 to: 251

1	TCACAGATCTGCGAGCATTCACCGCGAT...AGGACCAATCCCTCT	47
2	:     :     :     :     :     :     :     :     :	
113	TYGJYASpLeuGlnuAgyrPrAlaHisInsIserArgcInGlnIleProLe	129
48	GGGTATTAAGCAATCTATTCATACCCGTCGCGGCGCTTCGTTATCCAGCG	97
129	ucgIyLeuInAlaLeuThrHISgLYIleSerPhePheArgSerGlyJyA	146
98	GCAGCACCCGGGGCCAAAGCCGTCGTCCTTTAAATCCCAATGATGATATTC	147
146	snAspAsnGluGluYsAlaIArgThrLeuIleValIleIleGlnMetAla	162
148	TCGAGGGCGGGATATTCATCCCATCTTTGGAGGGCGCCCAATACAT	197
163	AlaAlaAlaAlaAlaArgPheArgYrIleIserAsnArgYrAlaYsAlaSerIle	179
198	TAAACGGGGAGACATCTCTCCGACATATACATCTCCAGCTCGACAGA	247
179	elcInIrrgLYrAlaPheInProAspAlaAlaMetIleIserLeuGlnA	196
248	CTAGTGTGGGGCCAAACATCCACGCAAGTCCAGCAGCTACGGATGGCTT	297
196	snAsnIrrgPAspAsnLeuSerAlaGlyValGlnGluSerValGlnAspHr	212
298	TTTAAATACCCATTTCGTTGGTGGGTATATTCACCGGTACTCTGTGACGTT	347
213	PheProAsnGln...ValThrLe	219
348	GAGCAATCTTCCGAC...G	364
219	YhrAsnIleIrrgLeuInuProValIleValAspSerLeuSerHisProt	236
365	TGATCGCCAGCTTAGCGATCATGTTGTTGTATGT	399
366	hcvAlaValaLeuAlaLeuMetLeuPheValGlyS	247

seq\_name: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:US-07-787-567A-1

```

1 seq_documentation_block:
2 /
3 / Sequence 1, Application US/07781567A
4 /
5 / GENERAL INFORMATION:
6 /
7 / APPLICANT: Better, Marc D.
8 /
9 / APPLICANT: Carroll, Steve F.
10 /
11 / APPLICANT: Lane, Julie A.
12 /
13 / TITLE OF INVENTION: Materials Comprising and Methods of
14 /
15 / TITLE OF INVENTION: Preparation and use for Ribosome-Inactivating Proteins
16 /
17 / NUMBER OF SEQUENCES: 36
18 /
19 / CORRESPONDENCE ADDRESS:
20 /
21 / ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
22 /
23 / ADDRESSEE: Ricknell
24 /
25 / STREET: Two First National Plaza, 20 South Clark
26 /
27 / STREET: Street
28 /
29 / CITY: Chicago
30 /
31 / STATE: Illinois
32 /
33 / COUNTRY: USA
34 /
35 / ZIP: 60603
36 /
37 / COMPUTER READABLE FORM:
38 /
39 / MEDIUM TYPE: Floppy disk
40 /
41 / COMPUTER: IBM PC compatible
42 /
43 / OPERATING SYSTEM: PC-DOS/MS-DOS
44 /
45 / SOFTWARE: PatentIn Release #1.0 Version #1.25

```

```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US-07/787,567A
3 FILING DATE: 199111104
4 CLASSIFICATION: 536
5 ATTORNEY/AGENT INFORMATION:
6 NAME: Gubler, Lewis S.
7 REGISTRATION NUMBER: 30,060
8 REFERENCE/DOCKET NUMBER: 27129/30545
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: (312) 346-5750
11 TELEFAX: 25-3856
12 INVENTOR FOR SEQ. ID. NO. 1:
13 SOURCE CHARACTERISTICS:
14 LENGTH: 267 amino acids
15 TYPE: AMINO ACID
16 TOPOLOGY: linear
17 MOLECULE TYPE: protein
18
19 US-07-787-567A-1

```

```
alignment_scores:
  Quality: 217.00
  Ratio: 2.309
  Percent Similarity: 65.734
  Length: 143
  Gaps: 2
  Percent Identity: 39.860
```

```
alignment_block:
```

Align seg 1/1 to: US-07-787-567A-1 from: 1 to: 267

[illegible]

```
seq_name: /cgn2_6/plotdata2/ppa/US0804_COMB.pep:US-08-406-801-33
seq_documentation_block:
; Sequence 34, Application US/08/406801
; GENERAL INFORMATION:
APPLICANT: Wright, Andrew F.
```

```

? APPLICANT: Blakey, David C.
? APPLICANT: Filton, John E.
? APPLICANT: Lindholm, Jelf
? APPLICANT: Lind, Peter
? APPLICANT: Holmgren, Jan
? TITLE OF INVENTION: Conjugates
? NUMBER OF SEQUENCES: 41
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Cushman, Darby & Cushman
? STREET: 1100 New York Avenue, N.W.
? CITY: Washington
? STATE: D.C.
? COUNTRY: U.S.A.
? ZIP: 20005-3918
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/406,801
? FILING DATE:
? CLASSIFICATION: 530
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 07/908,269
? FILING DATE: 02-JUL-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Kokulis, Paul N.
? REGISTRATION NUMBER: 16,773
? REFERENCE/DOCKET NUMBER: PNK/3893/96491/MLW
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-861-3000
? TELEFAX: 202-822-0944
? TELEX: 6714627 CUSH
? INFORMATION FOR SEQ ID NO: 34:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 267 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-406-801-34

```

```

alignment_scores:
  Quality: 217.00      Length: 143
  Ratio: 2.309         Gaps: 2
  Percent Similarity: 65.734   Percent Identity: 39.860

```

alignment\_block:

US-09-627-165B-15 x US-08-406-801-34 ..

Align seg 1/1 to: US-08-406-801-34 from: 1 to: 267

```

1 TACACAGATCTGAGGATACGCGGTCAT...AGGACGACATCCCTCT 47
||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:
123 TyrAspArgLeuGlnLeuAlaGlyAsnLeuArgGluAsnIleGluLe 139
48 GGGTATAGAGAACTCATTCATCCGTCGCGCGCTTCGTTAT..... 90
||| ||| :|||:|||||:|||||:|||||:|||||:|||||:
139 uGlyAsnGlyProLeuGlnLeuAlaIleSerAlaLeuTyrTyrSer 156
91 ..CCAGCGGCGACGACCGCGGCCCAAGCTGTTCCCTTATATCCGTCAT 138
||| :||| :|||:|||||:|||||:|||||:|||||:|||||:
156 hrcGlyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
139 CAGATGATCTCGAGCGCGGAGATTCATCCATCCCTTTGGAGGGCTCG 188
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 GlnMetIleSerGlnAlaAlaArgPheGlnTyrIleGluGluMetArg 189
189 CCAATATACATTAAACAGCGGAGAGTCATTTCTCCCGCATGATGATGCTCG 238
| :|||: :|||: :|||: :|||: :|||: :|||: :|||:
189 gThrArgIleArgTyrAsnArgArgSerAlaProAspProSerValIle 206

```

```

239 AGCTGAGAACTAGTGGGCGCCCAACAATCCAGCGCACTCAGAGTCTACG 288
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 hrcLeuAsnSerTrpGlyArgLeuSerThrAlaIleGlnIleAsn 222
289 GATGCGGCTTTTAAATACCATTTTCGTTGGGTATATCCACCGGTAAC 338
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySer 239
339 CGTACGATTGAGCAATGTTCCGCGACGTGATCCGACCTAGGACATCATG 388
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
239 sPheSerValTyrAspValSerIleLeuIleProIleIleAlaLeuMet 256
389 TGTTTGATGTAGGACGACCATCTTCC 417
:|||||:|||||:|||||:|||||:|||||:|||||:
256 alTyrArgCysAlaProProProSerSer 265

seq_name: /cgn2_6/ptodata/2/paa/US095_COMB.pep:US-09-538-873-1

seq_documentation_block:
? Sequence 1, Application US/09538873
? GENERAL INFORMATION:
? APPLICANT: VITETTA, ELLEN S.
? APPLICANT: GHETTE, VICTOR F.
? APPLICANT: SMALISHAW, JOAN
? APPLICANT: BALUNA, ROXANA G.
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
? FILE REFERENCE: US09538873
? CURRENT APPLICATION NUMBER: US/09/538,873
? EARLIER FILING DATE: 2000-03-30
? EARLIER APPLICATION NUMBER: 60/126,826
? NUMBER OF SEQ ID NOS: 19
? SOFTWARE: Patent In Ver. 2.1
? SEQ ID NO 1
? LENGTH: 267
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Synthetic
? OTHER INFORMATION: Peptide
? US-09-538-873-1

```

```

alignment_scores:
  Quality: 217.00      Length: 143
  Ratio: 2.309         Gaps: 2
  Percent Similarity: 65.734   Percent Identity: 39.860

```

alignment\_block:

US-09-627-165B-15 x US-09-538-873-1 ..

Align seg 1/1 to: US-09-538-873-1 from: 1 to: 267

```

1 TACACAGATCTGAGGATACGCGGTCAT...AGGACGACATCCCTCT 47
||| |||||:|||||:|||||:|||||:|||||:|||||:
123 TyrAspArgLeuGlnLeuAlaGlyAsnLeuArgGluAsnIleGluLe 139
48 GGGTATAGAGAACTCATTCATCCGTCGCGCGCTTCGTTAT..... 90
||| ||| :|||:|||||:|||||:|||||:|||||:|||||:
139 uGlyAsnGlyProLeuGlnLeuAlaIleSerAlaLeuTyrTyrSer 156
91 ..CCAGCGGCGACGACCGCGGCCCAAGCTGTTCCCTTATATCCGTCAT 138
||| :||| :|||:|||||:|||||:|||||:|||||:|||||:
156 hrcGlyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
139 CAGATGATCTCGAGCGCGGAGATTCATCCATCCCTTTGGAGGGCTCG 188
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 GlnMetIleSerGlnAlaAlaArgPheGlnTyrIleGluGluMetArg 189
189 CCAATATACATTAAACAGCGGAGAGTCATTTCTCCCGCATGATGATGCTCG 238
| :|||: :|||: :|||: :|||: :|||: :|||: :|||:
189 gThrArgIleArgTyrAsnArgArgSerAlaProAspProSerValIle 206

```

```

239 AGCTGAGACTAGTGGGGCCAAATCCAGCAGATCCAGAGTCTACG 288
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 hrleuglnasnserrtpgllyargluserthrAlaIleGlnIuserran 222
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
289 GATGGCGTTTAAATACCATTTGGGTTGGTATATCCAGCGGTACTG 338
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerIy 239
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
339 CGTGAAGTGAAGCATGTTTCGACGCGTGAATGCCAGCTTAGCATCATGT 388
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
239 sPheserValTyrAspValSerIleuIleuIleIleAlaIleuMetV 256
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
389 TGTGTGTATGTAAGGACCGACCATCTTC 417
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
256 alyrArgCysAlaProProSerSer 265
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:

seq_name: /cgn2_6/ptodata/2/paa/US096_COMB.pep:US-09-610-838-1

seq_documentation_block:
/ Sequence 1, Application US/09610838
/ GENERAL INFORMATION:
/ APPLICANT: Better, Marc D.
/ APPLICANT: Carroll, Stephen F.
/ APPLICANT: Studlika, Gary M.
/ TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
/ TITLE OF INVENTION: Proteins
/ NUMBER OF SEQUENCES: 173
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: McAndrews, Held & Malloy, Ltd.
/ STREET: 500 West Madison Street, 34th floor
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60661
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/610,838
/ FILING DATE: 06-JUL-2000
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/09/136,389
/ FILING DATE: 18-AUG-1998
/ APPLICATION NUMBER: 08/646,360
/ FILING DATE: 13-MAY-1996
/ APPLICATION NUMBER: PCT/US94/05348
/ FILING DATE: 12-MAY-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/064,691
/ FILING DATE: 12-MAY-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/988,430
/ FILING DATE: 09-DEC-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/901,707
/ FILING DATE: 19-JUN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/787,567
/ FILING DATE: 04-NOV-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McNicholas, Janet M.
/ REGISTRATION NUMBER: 32,918
/ REFERENCE/DOCKET NUMBER: 200-70, P4
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/707-8889
/ TELEFAX: 312/707-9155
/ TELEX: 650 388-1248
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-610-838-1

alignment_scores:
Quality: 217.00 Length: 143
Ratio: 2.309 Gaps: 2
Percent Similarity: 65.734 Percent Identity: 39.860

alignment_block:
US-09-627-165b-15 x US-09-610-838-1 ..

Align seg 1/1 to: US-09-610-838-1 from: 1 to: 267

1 TACACAGATCTGGAGCGATACCGCGTCTAT...AGGACCGAGATCCCTCT 47
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
123 TyrAspArgLeuGlnGlnLeuAlaIleGlnLeuArgGlnSerIleGluLe 139
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
48 GGGTATAGAGCAACTCATTCATCCGCTCGCGCCCTCGTTAT..... 90
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 uGlyAsnGlyProLeuGlnGlnAlaIleSerAlaIleuTyrTyrIleTyr 156
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 ..CCAGGCGGACGACCGCGGCCCAAGCTCGTCCCTTATATATCTCAT 138
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156 hrGlyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 CAGATGATCTCCGAGCGCGCGAGATTCATCCATCTTTGAGAGGCTCG 188
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 GlnMetIleSerGlnAlaAlaArgPheGlnIleTyrIleGlnGlnMetAr 189
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
189 CCAATATCATTAACAGCGGCGAGTATCTTCCGACATGATGATGCTCG 238
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 gThrArgIleArgTyrAsnArgArgSerAlaProAspProSerValIleT 206
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
239 AGCTGAGACTAGTGGGGCCAAATCCAGCAGATCCAGTCTAGC 288
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 hrleuglnasnserrtpgllyargluserthrAlaIleGlnIuserran 222
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
289 GATGGCGTTTAAATACCATTTGGGTTGGTATATCCAGCGGTACTG 338
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerIy 239
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
339 CGTGAAGTGAAGCATGTTTCGACGCGTGAATGCCAGCTTAGCATCATGT 388
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
239 sPheserValTyrAspValSerIleuIleuIleIleAlaIleuMetV 256
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
389 TGTGTGTATGTAAGGACCGACCATCTTC 417
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
256 alyrArgCysAlaProProSerSer 265
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:

seq_name: /cgn2_6/ptodata/2/paa/US096_COMB.pep:US-09-668-419-1

seq_documentation_block:
/ Sequence 1, Application US/09668419
/ GENERAL INFORMATION:
/ APPLICANT: VITETTA, ELLEN S.
/ APPLICANT: GHETTE, VICTOR F.
/ APPLICANT: SMALLSHAW, JOAN E.
/ APPLICANT: BALDUNA, ROXANA G.
/ TITLE OF INVENTION: RICIN A CHAIN MUTANTS LACKING ENZYMATIC ACTIVITY AS VACCINES
/ FILE REFERENCE: UTSD:72305
/ CURRENT APPLICATION NUMBER: US/09/668,419
/ CURRENT FILING DATE: 2000-09-22
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 267
/ TYPE: PRT
/ ORGANISM: Artificial Sequence

```

```

: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
: OTHER INFORMATION: Peptide
US-09-668-419-1

```

```
alignment_scores:
  Quality: 217.00
  Ratio: 2.309
  Percent Similarity: 65.734
  Length: 143
  Gaps: 2
  Percent Identity: 39.860
```

```
alignment_block:
```

Align seg 1/1 to: US-09-668-419-1 from: 1 to: 267

```

    1 TACAGATGTCGGAGCGCATTTACGCCGCATC... AGGACCAAGATCCCTT 47
      ||| |||||::: |||||::: |||||:::
   123 TyrNspArgLeuGlInL euAlaIGlyAsnLeuAr gLuAnIleGuLe 139
     48 GGGATTAGAAGAACTCAATTCAATCCGTCTGGCGCTTGCTTAT..... 90
       |||| | || :|||||::::::::::|||
   139 uGIysnGLyProLeuglUGluAIalIEserAlaleuTYrTyrisert 156
     91 ..CCAAGCGGAGACACC GGSCGCCAAGCTGGTCCCTTAATAACCTCAT 138
        ||| ::| :||| ::||| |||||:::||||| |||
   156 hncLYgLyThrcInleuProthrleuAlar gSerPheIIeICysile 172
     139 CAGATGATCTCCGAGGCCGCGAGATTCACATCCCATCTTTGGAGGCTCG 188
       173 GlmETIlIESerGUAlAlAlARphEGlnTYrllegLUglYguNEtar 189
         189 CCAAATACATTAACAGCGGAGCATCTTCTCCCCGACATGAACAATGCTCG 238
           189 gThrArglle rgyrrasnaatAg serAlaprosP roseVallet 206
             239 ACCTGAGACTTAGTTGGGCGCACAAATCCACAGCAAGTCCAGACTTCAG 288
               ||||| ::|||::: |||||::: |||||:::
            206 htLEuGluASerTErrPol yArgLeuSErrhralIIEglngIUserAm 222
              289 GATGAGCGTTTTAATAACCCATCTTCGSTGGGTATATCCAACCGCTAATT 358
                ::|||:||| :|||||::::::::::||| :|||||:::
          223 glncIglaAPheaIasePrOIleIngUalnArgIarscnIgeSery 239
            339 CGTGACGTTAGAGATGTTCCGCAGCATGATCGCCAGCTTAgscATCATGT 388
              ::||| ::||| ::||| ::||| ::||| ::|||
          239 spHeSe rValtyraspalSer Ile leuIleproIle lAlalauewEv 256
            389 TGTTTGATGNAGGACCGACCATCTCC 417
              ::||| ||| ||||| |||||
          256 alTYra fCYsa la p robe Pro se rder 265
```



OM of: US-09-627-165B-15 to: Pending\_Patents\_AA\_New.\* out\_format : pfs

Date: Jul 30, 2002 4:07 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame2p model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_SPOOL/US09627165/unal\_30072002\_151139\_3273/app\_query.fasta.1.1012  
-DB=Pending\_Patents\_AA\_New -OFMT=fastan -SUFFIX=rapn  
-GAPOP=12.000 -GAPOP=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
-LOOPEXT=0.000 -GAPOP=6.000 -FEAPEXT=7.000 -XGAPOP=10.000  
-XGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=biosum62 -TRANS=human40.cdi -LIST=45 -DLOCALIGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=2000000000  
-USER=US09627165 -CGNL 1.45 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPRX -WAIT -THREADS=1

Search information block:

Query: US-09-627-165B-15  
Query length: 522  
Database: Pending\_Patents\_AA\_New.\*  
Database sequences: 447039  
Database length: 145595129  
Search time (sec): 104.820000

score\_list:

Sequence	Strd Orig	zScore	EScore	len	Documentation	
/cgn2_6/ptodata/1/paa/US09_NEW.COMB.pcp:US-09-601-667B-37		621.00	1189.04	2.3e-58		2
/cgn2_6/ptodata/1/paa/US09_NEW.COMB.pcp:US-09-601-667B-39		604.50	1157.04	1.4e-56		2
/cgn2_6/ptodata/1/paa/US09_NEW.COMB.pcp:US-09-601-667B-42		599.00	1146.28	5.5e-56		2
/cgn2_6/ptodata/1/paa/US09_NEW.COMB.pcp:US-09-601-667B-41		599.00	1146.24	5.5e-56		2
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-083-336A-1		259.50	478.60	3.8e-19		57
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-1		217.00	403.97	1.2e-14		267
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-09-726-651A-4		156.50	285.67	4.2e-08		28
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-6		153.50	281.46	8.5e-08		247
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-9		152.50	268.95	1.1e-07		261
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-10		147.00	268.67	4.3e-07		24
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-11		147.00	268.67	4.3e-07		24
/cgn2_6/ptodata/1/paa/US09_NEW.COMB.pcp:US-09-663-826-7		145.50	265.88	6.2e-07		248
/cgn2_6/ptodata/1/paa/US09_NEW.COMB.pcp:US-10-127-890-4		141.00	256.54	1.9e-06		263
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-2		136.00	245.37	8.5e-06		25
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-99		135.00	245.37	8.5e-06		25
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-100		135.00	245.37	8.5e-06		25
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-101		135.00	245.37	8.5e-06		25
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-106		135.00	245.37	8.5e-06		25
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-107		135.00	245.37	8.5e-06		25
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-108		135.00	245.37	8.5e-06		25
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-109		135.00	245.37	8.5e-06		25
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-110		135.00	245.37	8.5e-06		25
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-110		135.00	245.37	8.5e-06		25
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-104		130.00	235.68	3.0e-05		2
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-103		124.00	223.68	0.0001		25
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-10		116.00	208.59	0.0010		248
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-127-890-9		105.50	191.22	0.0120		18
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-083-336A-9		105.50	191.05	0.0121		18
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-083-336A-8		105.50	191.05	0.0121		18
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-083-336A-7		105.50	190.94	0.0121		18
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-083-336A-11		105.50	190.52	0.0123		19
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-083-336A-3		105.50	190.52	0.0123		19
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-083-336A-5		105.50	190.42	0.0123		19
/cgn2_6/ptodata/1/paa/US10_NEW.COMB.pcp:US-10-083-336A-10		81.50	135.06	6.37		2

seq\_name: /cgn2\_6/ptodata/1/paa/US09\_NEW.COMB.pcp:US-09-601-667B-37

seq\_documentation\_block:

Sequence 37, Application US/09601667B  
GENERAL INFORMATION:  
APPLICANT: Morits, Peter  
APPLICANT: Stiefel, Thomas  
APPLICANT: Voelter, Wolfgang  
APPLICANT: Welters, Peter  
TITLE OF INVENTION: Recombinant Mistletoe Lectins  
FILE REFERENCE: 29841/36636  
CURRENT APPLICATION NUMBER: US/09/601,667B  
CURRENT FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: PCT/EP99/00696  
PRIOR FILING DATE: 1999-02-03  
PRIOR APPLICATION NUMBER: D 198 04 210.8  
NUMBER OF SEQ ID NOS: 41  
SEQ ID NO 37  
LENGTH: 254  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: mistletoe lectin A1  
US-09-601-667B-37

alignment\_scores:

Quality: 621.00 Length: 140  
Ratio: 4.634 Gaps: 0  
Percent Similarity: 95.714 Percent Identity: 85.714

alignment\_block:

US-09-627-165B-15 x US-09-601-667B-37 ..  
Align seg 1/1 to: US-09-601-667B-37 from: 1 to: 254

1	TACACAGATCTGAGGATACCGCGCTATAGGACCAAGATCCCTCGG	50
111		
115	TYRPRDAPLEUCILNARYTYRGLYISATRYGASPCINLLEPROLEUGL	131
111		
51	TATAGAGAACATTCATTCGCTCGCGCTGATTCAGAGGCA	100
111		
131	YILEAPRGINLEILLGLSERVALTHRLALEUVRGHPROGLYGLYS	148
111		
101	GCACCCGCCCAAGCTCGTCCCTTATATCCATTCAGATCTCC	150
111		
148	ERTHRARGHRLGALALARGSERILEULEILLGLMETLLESER	164
111		
151	GAGCCGCCAGATTCATCCCATCTTTTGAGAGCCTCCGCAATACATTA	200
111		
165	GLALALALARGPHEASNPROLLEUTRPARGYTYRATGINTTYLLEAS	181
111		
201	CAGCGGGGATCTTCTCCGACATGTACATGCTCGAGTGGAGACTA	250
111		
181	NSEGLYALASERPHLEUPROASPYALTYRMETLEUGLLEUJLHNS	198
111		
251	GTTGGGCCAACATCCAGCAGATCCAGAGTTCACGATGCGCTTTT	300
111		
198	ERTPRGLYGLINLSERTHGLNVALGHLNLSERTHRSPLGYALPHE	214
111		
301	AATTAACCATTTGGTGGTATATACACCGGTAATCTGTACGTTGAG	350
111		
215	ASNSNPROLLEARGLEUALALIEPROPOGLYASNPHEVALTHLEUWH	231
111		
351	CAATGTTGCGAGTGTATCCAGCTTAGCGATGATGTTGTATGTA	400
111		
231	RASNLVALRGASPYALLLEALASERLEUALILEMETLEUPHEVALCYG	248
111		
401	GGGACCGACATCTTCTCC	420
111		
248	LYGLIARGPROSERSER	254

```
seq.name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-39
seq_documentation_block:
; Sequence 39, Application US/09601667B
; GENERAL INFORMATION:
; APPLICANT: Morris, Peter
; APPLICANT: Stiefel, Thomas
; APPLICANT: Voelter, Wolfgang
; APPLICANT: Welters, Peter
; TITLE OF INVENTION: Recombinant Mistletoe Lectins
; FILE REFERENCE: 29841/36636
; CURRENT APPLICATION NUMBER: US/09/601,667B
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: PCT/EP99/00696
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: D 198 04 210.8
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 41
; SEQ ID NO 39
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mistletoe lectin A1 (matched)
US-09-601-667B-39

alignment_scores:
Quality: 604.50 Length: 140
Ratio: 4.545 Gaps: 1
Percent Similarity: 95.000 Percent Identity: 85.000

alignment_block:
US-09-627-165B-15 x US-09-601-667B-39 ..

Align seg 1/1 to: US-09-601-667B-39 from: 1 to: 253

1 TACACAGATCTGAGGATACGGCGGATACGAGGACCAATCCCTGTGGG 50
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115 TyrProAspLeuGlnAlaArgSerIleLeuIleGlnMetIleSer 130
51 TATAGAGGAATCATTCATCGCTCGGCGCTTATACGAGCGGCGA 100
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
130 ylleAspGlnleuIleGlnSerValThrAlaLeuArgPheProGlyGlyS 147
101 GCACCGCGGCCCAAGCTCGTCCCTTATATCCTCATTCAGATGATCTCC 150
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
147 errThrArgThrGlnAlaArgSerIleLeuIleGlnMetIleSer 163
151 GAGCGCGGATTCATCCATCTTTGGAGGCGTCCGATACATTA 200
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
164 GlnAlaAlaArgPheAsnProIleLeuThrArgIleGlnMetIleS 180
201 CACCGCGGAGTCAATTCCTCCGACATGATGCTCGAGCTGAGACTA 250
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
180 nSerGlnAlaSerPheLeuProAspValTyrMetLeuGlnLeuGln 197
251 GTTGGGGCCCAATCCAGCAAGTCCAGCATTCAGAGTGGCTTTT 300
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
197 errIlePylGlnGlnSerThrGlnValGlnHisSerThrAspIleVal 213
301 AATATCCATTTGCGTGGGTATATCCACCGGATCACTTCGAGCTGAG 350
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
214 AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValThrLeu 230
351 CAATGTCGCGAGCGATGATCCAGCTTACCGCATTCATGTTGTATGTA 400
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
230 rAsnValAlaArgAspValIleAlaSerLeuAlaIleMetLeuPheVal 247
401 GGGACCGACCATCTTCCTCC 420
||| ||||| ||||| |||||
247 LylGlnArgProSerSerSer 253
```

```
seq.name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-2
seq_documentation_block:
; Sequence 2, Application US/09601667B
; GENERAL INFORMATION:
; APPLICANT: Morris, Peter
; APPLICANT: Stiefel, Thomas
; APPLICANT: Voelter, Wolfgang
; APPLICANT: Welters, Peter
; TITLE OF INVENTION: Recombinant Mistletoe Lectins
; FILE REFERENCE: 29841/36636
; CURRENT APPLICATION NUMBER: US/09/601,667B
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: PCT/EP99/00696
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: D 198 04 210.8
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 41
; SEQ ID NO 2
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: M1A-chain
; NAME/KEY: SITE
LOCATION: 15
OTHER INFORMATION: product= "Xaa is Asp or Glu"
FEATURE:
NAME/KEY: SITE
LOCATION: 63
OTHER INFORMATION: product= "Xaa is Gly or Gln"
OTHER INFORMATION: /label= Xaa2
FEATURE:
NAME/KEY: SITE
LOCATION: 66
OTHER INFORMATION: product= "Xaa is Ile or Val"
OTHER INFORMATION: /label= Xaa3
FEATURE:
NAME/KEY: SITE
LOCATION: 75
OTHER INFORMATION: product= "Xaa is Leu or Ala"
OTHER INFORMATION: /label= Xaa4
FEATURE:
NAME/KEY: SITE
LOCATION: 107
OTHER INFORMATION: product= "Xaa is missing"
OTHER INFORMATION: /label= Xaa5
FEATURE:
NAME/KEY: SITE
LOCATION: 113
OTHER INFORMATION: product= "Xaa is Asn or Thr"
OTHER INFORMATION: /label= Xaa6
FEATURE:
NAME/KEY: SITE
LOCATION: 117
OTHER INFORMATION: product= "Xaa is Pro or Thr"
OTHER INFORMATION: /label= Xaa7
FEATURE:
NAME/KEY: SITE
LOCATION: 141
OTHER INFORMATION: product= "Xaa is Ser or Thr"
OTHER INFORMATION: /label= Xaa9
FEATURE:
NAME/KEY: SITE
LOCATION: 145
```



```
OTHER INFORMATION: product= "Xaa is Phe or Tyr"
OTHER INFORMATION: /label= Xaa10
FEATURE:
NAME/KEY: SITE
LOCATION: 152
OTHER INFORMATION: product= "Xaa is Thr or Ala"
OTHER INFORMATION: /label= Xaa11
FEATURE:
NAME/KEY: SITE
LOCATION: 177
OTHER INFORMATION: product= "Xaa is Ala or Tyr"
OTHER INFORMATION: /label= Xaa12
FEATURE:
NAME/KEY: SITE
LOCATION: 180
OTHER INFORMATION: product= "Xaa is Tyr or Asp"
OTHER INFORMATION: /label= Xaa13
FEATURE:
NAME/KEY: SITE
LOCATION: 185
OTHER INFORMATION: product= "Xaa is Ala or Glu"
OTHER INFORMATION: /label= Xaa14
FEATURE:
NAME/KEY: SITE
LOCATION: 191
OTHER INFORMATION: product= "Xaa is Val or Met"
OTHER INFORMATION: /label= Xaa15
FEATURE:
NAME/KEY: SITE
LOCATION: 219
OTHER INFORMATION: product= "Xaa is Ile or Phe"
OTHER INFORMATION: /label= Xaa16
FEATURE:
NAME/KEY: SITE
LOCATION: 224
OTHER INFORMATION: product= "Xaa is Pro or Ser"
OTHER INFORMATION: /label= Xaa17
FEATURE:
NAME/KEY: SITE
LOCATION: 225
OTHER INFORMATION: product= "Xaa is Pro or Thr"
OTHER INFORMATION: /label= Xaa18
FEATURE:
NAME/KEY: SITE
LOCATION: 232
OTHER INFORMATION: product= "Xaa is Thr or Ser"
OTHER INFORMATION: /label= Xaa19
FEATURE:
NAME/KEY: SITE
LOCATION: 236
OTHER INFORMATION: product= "Xaa is Asp or Ser"
OTHER INFORMATION: /label= Xaa20
US-09-601-667B-2

alignment_scores:
Quality: 599.00      Length: 140
Ratio: 4.538         Gaps: 0
Percent Similarity: 94.286      Percent Identity: 84.286

alignment_block:
US-09-627-165B-15 x US-09-601-667B-2 ..
Align seg 1/1 to: US-09-601-667B-2 from: 1 to: 255

1 TACACAGATCTGGAGGATACGCCGGTTCATAGGACAGATCCCTCTGGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 Tyr***AspLeuGlnArgTyrIleGlyHisArgAspGlnIleProLeuG1 132
51 TATAGAGGAACTCATTCATCCGCTCGCGCGCTTCCTTATCCAGGCGGCA 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 ylle**GlnIleuIleGlnInsVal**AlaIleuArg**ProGlyGlys 149
```

```
101 GCACCGGCGCCAGCTGCTCCCTTAATTCCTCATTCAGATGATCTCC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149 erThrArg***GlnIleArgSerIleIleuIleGlnMetIleSer 165
151 GAGCGCGGAGATTCATTCATCTTTTGGAGGCGCTCCGCAATACATTA 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 GlnAlaIleArgPheAsnProIleIleuIleTrpArg***ArgGln***IleAs 182
201 CAGCGGGGAGTCATCTTCGCCAGATGACATGCTGAGCTGGAGACTA 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 nserGly***SerPheLeuProAsp***TyrMetLeuGlnIleuGlnThrS 199
251 GTTGGGCGCCAGATTCACAGCAAGTCCAGAGTCTAGAGATGGGCTTTT 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 erTrpGlnGlnInsSerThrGlnVal3IleHisSerThrAspGlyValPhe 215
301 AATRACCATTCCTCGTGGTATATCCACCGGTAATCTCGATGGATGAG 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 AsnAsnPro***ArgLeuAlaIle****GlyAsnPheValThrLeu** 232
351 CAATGTCGCGAGCTGATCGCCAGCTTACCGATCATGTTGTTGTATGTA 400
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 *AsnValArg***ValIleIleAserIleuAlaIleMetLeuPheValCysG 249
401 GCGACCGACCATCTCTCTCC 420
:::|||||:|||||:|||||
249 IyGlnArgProSerSerSer 255

seq_name: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-41
seq_documentation_block:
; Sequence 41, Application US/0960167B
; GENERAL INFORMATION:
; APPLICANT: Morris, Peter
; APPLICANT: Stiefel, Thomas
; APPLICANT: Voelter, Peter
; APPLICANT: Walters, Peter
; TITLE OF INVENTION: Recombinant Mistletoe Lectins
; FILE REFERENCE: 29841/36636
; CURRENT FILING DATE: US/09/601,667B
; PRIOR APPLICATION NUMBER: PCT/EP99/00696
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: D 198 04 210.8
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 41
; SEQ ID NO 41
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: MUA-chain
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (15)..(15)
; OTHER INFORMATION: Xaa is Asp or Glu
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (66)..(66)
; OTHER INFORMATION: Xaa is Ile or Val
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (75)..(75)
; OTHER INFORMATION: Xaa is Leu or Ala
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (114)..(114)
; OTHER INFORMATION: Xaa is Asn or Thr
; FEATURE:
```

```

NAME/KEY: SITE
LOCATION: (118)..(118)
OTHER INFORMATION: xaa is Pro or Thr
FEATURE:
NAME/KEY: SITE
LOCATION: (135)..(135)
OTHER INFORMATION: xaa is Asp or Glu
FEATURE:
NAME/KEY: SITE
LOCATION: (142)..(142)
OTHER INFORMATION: xaa is Ser or Thr
FEATURE:
NAME/KEY: SITE
LOCATION: (146)..(146)
OTHER INFORMATION: xaa is Phe or Tyr
FEATURE:
NAME/KEY: SITE
LOCATION: (153)..(153)
OTHER INFORMATION: xaa is Thr or Ala
FEATURE:
NAME/KEY: SITE
LOCATION: (178)..(178)
OTHER INFORMATION: xaa is Ala or Tyr
FEATURE:
NAME/KEY: SITE
LOCATION: (181)..(181)
OTHER INFORMATION: xaa is Tyr or Asp
FEATURE:
NAME/KEY: SITE
LOCATION: (186)..(186)
OTHER INFORMATION: xaa is Ala or Glu
FEATURE:
NAME/KEY: SITE
LOCATION: (192)..(192)
OTHER INFORMATION: xaa is Val or Met
FEATURE:
NAME/KEY: SITE
LOCATION: (220)..(220)
OTHER INFORMATION: xaa is Ile or Phe
FEATURE:
NAME/KEY: SITE
LOCATION: (225)..(225)
OTHER INFORMATION: xaa is Pro or Ser
FEATURE:
NAME/KEY: SITE
LOCATION: (226)..(226)
OTHER INFORMATION: xaa is Pro or Thr
FEATURE:
NAME/KEY: SITE
LOCATION: (233)..(233)
OTHER INFORMATION: xaa is Thr or Ser
FEATURE:
NAME/KEY: SITE
LOCATION: (237)..(237)
OTHER INFORMATION: xaa is Asp or Ser
US-09-601-667B-41

```

```

alignment_scores:
Quality: 599.00      Length: 140
Ratio: 4.538        Gaps: 0
Percent Similarity: 94.286      Percent Identity: 84.286

```

alignment\_block:  
US-09-627-165B-15 x US-09-601-667B-41 ..

Align seg 1/1 to: US-09-601-667B-41 from: 1 to: 256

```

1 TACACGATCTGAGAGATACGCGGTCAATAGGAGACAGATCCCTGAG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
117 Tyr***AspLeuGluArgTyrAlaGlyHisArgAspGlnIleProIleu 133
51 TATAGAGAGACATTCATTCGCTCTCGGCGCTTGTATCCAGCGGCA 100

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
133 Ylle***GlnIleuIleGlnSerVal***AlaLeuArg***ProGlyGly 150
101 GCACCCGGGCCCCAAGCTCGTTCCTTATATATCCATTCAGATGATCC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
150 eXrHrArg***GlnAlaArgSerIleIleuIleGlnMeIleSer 166
151 GAGCCGGAGATTCATTCATCCATCTTGTAGAGGCTCCCAATCATTA 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 GluAlaIleArgPheAsnProIleuLeuTyrParg***ArgGln***Ile 183
201 CAGCGGGAGATCTTCTCCGACATTCATGCTGACGCTGAGACTA 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
183 nserGly***SerPheLeuProAsp***TyrMeLeuGlnLeuGln 200
251 GTTGGGCCCAACATTCAGCGCAAGTCCAGAGTCTAGGATGGCTTTT 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
200 eXrTpGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 216
301 AATAACCATTTGCGTTGGGATATCCACCGGTAACCTTCGTGAGTTGAG 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
217 AsnAsnPro***ArgLeuAlaIle***GlyAsnPheValThrLeu*** 233
351 CAATGTCGCGACGTGATCCGACGCTTACGATCATGTGTTGTATGTA 400
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
233 *AsnValArg***ValIleAlaSerLeuAlaIleMeLeuPheValGly 250
401 GCGACGACCATCTTCCTCC 420
:::|||||:|||||:|||||
250 YgluArgProSerSerSer 256

```

seq\_name: /cgn2\_6/prodata/1/paa/US10\_NEW\_COMB.pcp:US-10-083-336A-1

```

seq_documentation_block:
; Sequence 1, Application US/10083336A
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Rizin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452050 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083.336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 576
; TYPE: prot
; ORGANISM: Rizinus communis
US-10-083-336A-1

```

```

alignment_scores:
Quality: 259.50      Length: 178
Ratio: 2.257        Gaps: 5
Percent Similarity: 64.607      Percent Identity: 41.011

```

alignment\_block:

US-09-627-165B-15 x US-10-083-336A-1 ..

Align seg 1/1 to: US-10-083-336A-1 from: 1 to: 576

```

1 TACACGATCTGAGAGATACCGCGGTCAATAGGAGACAGATCCCTCT 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
158 TyrAspArgLeuGlnGlnIleuAlaGlyAsnLeuArgGlnAsnIleGlu 174
48 GGGTATAGAGAGACATTCATTCATCCGCTCGGCGCTTGTAT..... 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
174 uGlyAsnGlyProLeuGlnGlnAlaIleSerAlaLeuTyrTyrIle 191
91 ..CCAGGCGGACGACCCGGGCCCAAGCTCGTTCCTTATATCTCAT 138
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

191 hrclyglythrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 207
139 CAGATGATCTCCGAGCCCGGAGATTCATCCATCTTTTGAGGCGCTCG 188
|||||
208 GlnMetIleSerGlnAlaAlaArgPheGlnIleIleGlnGlnIleMet 224
189 CCAATACATTACAGCGGGAGTCATTTCTCCGACATGTCATGCTCG 238
|||||
224 gthrrArgIleArgTyrAsnArgArgSerAlaIleProAspProSerValIle 241
239 AGCTGAGACTAGTTGGGGCCCAACATCCAGCAATGCCAGCTAGCTAG 288
|||||
241 hrlengGlnAsnSerTrpGlyArgLeuSerThrAlaIleGlnGlnIleAsn 257
289 GATGGCGTTTAAATACCATTCCTGGGTGGTATTCACCGGTAACTT 338
|||||
258 GlnGlyAlaPheAlaSerProIleGlnLeuGlnIleArgArgAsnGlySer 274
339 CGTACGCTTGAGCAATGTTCGCGAGCTGATCGCAGCTTACCATCAT 388
|||||
274 sPheSerValIleTyrAspValSerIleLeuIleProIleIleAlaLeuMet 291
389 TGTTGTATGTAGAGCAGCCGACCATCTCTCCGAGCTCGCTATTTGCG 438
|||||
291 alTyrArgCysAlaProProProSerSerGln.....PheSer 303
439 CTGGTCATATGACCGCTTGTGAAATATAGCGCGCGCTGACGATGTTAC 488
|||||
304 LeuLeuIleArgProValIleAlaProAsnPheAsnAla.....AspVal.. 317
489 CTCGACTGCTCCGAMCCGACCGCTGGCATCGTA 522
|||||
318 .CysMetAspProGlnProIleValArgIleVal 328

seq_name: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:US-10-127-890-1

seq_documentation_block:
; Sequence 1, Application US/10127890
; GENERAL INFORMATION:
; APPLICANT: Belter, Marc D.
; Studnika, Gary M.
; Carroll, Stephen F.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAnotews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/127,890
; FILING DATE: 23-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; APPLICATION NUMBER: US 07/787,567

```

```

; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70..P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-9155
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-127-890-1

```

```

alignment_scores:
Quality: 217.00 Length: 143
Ratio: 2.309 Gaps: 2
Percent Similarity: 65.734 Percent Identity: 39.860

```

```

alignment_block:
US-09-627-165B-15 x US-10-127-890-1 ..

```

```

Align seg 1/1 to: US-10-127-890-1 from: 1 to: 267

```

```

1 TACACAGATCTGAGCGATACGCCGTCAT...AGGACCATCTCTCT 47
|||||
123 TyrAspArgLeuGlnGlnLeuAlaGlyAsnLeuArgGlnAsnIleGlnIle 139
48 GGGTATAGAGCACTCATTCACCTCTCGCGCTCGTAT..... 90
|||||
139 uclYasnGlyProLeuGlnGlnAlaIleSerAlaLeuTyrTyrSer 156
91 ..CCAGCGCGCAGCACCCGCGCCAGCTCGTCCCTTAATTCAT 138
|||||
156 hrclyglythrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
139 CAGATGATCTCCGAGCGCGGATTCATCCATCTTTTGAGGCGCTCG 188
|||||
173 GlnMetIleSerGlnAlaAlaArgPheGlnIleIleGlnGlnIleMet 189
189 CCAATACATTACAGCGGGAGTCATTTCTCCGACATGTCATGCTCG 238
|||||
189 gthrrArgIleArgTyrAsnArgArgSerAlaIleProAspProSerValIle 206
239 AGCTGAGACTAGTTGGGGCCCAACATCCAGCAATGCCAGCTAGCTAG 288
|||||
206 hrlengGlnAsnSerTrpGlyArgLeuSerThrAlaIleGlnGlnIleAsn 222
289 GATGGCGTTTAAATACCATTCCTGGGTGGTATTCACCGGTAACTT 338
|||||
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnIleArgArgAsnGlySer 239
339 CGTACGCTTGAGCAATGTTCGCGAGCTGATCGCAGCTTACCATCAT 388
|||||
239 sPheSerValIleTyrAspValSerIleLeuIleProIleIleAlaLeuMet 256
389 TGTTGTATGTAGGAGCAGCATCTTCC 417
|||||
256 alTyrArgCysAlaProProProSerSer 265

```

```

seq_name: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:US-09-726-651A-4

seq_documentation_block:
; Sequence 4, Application US/0972651A
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: DANSON, William O.
; APPLICANT: GRANTHAM, GEORGE L.

```



alignment\_block:  
US-09-627-165b-15 x US-10-127-890-7 ..

Align seg 1/1 to: US-10-127-890-7 from: 1 to: 263

```

1 TACACAGATCTGAGCGATACGCCGCTCAT...AGGACACAGATCCCTCT 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
111 TyrgluarGleuglnleAlaAlaGlyLysProArGlnLysIleProIle 127
48 GGGTTATAGAGCAATTCATCCGCTCTGCGCGCTTCCTGATTCAGCGCG 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127 eGlyLeuProAlaLeuAspSerAlaIleSerThrLeuLeuHis.....T 142
98 GCAGACCGCGCGCCAGCTGCTCCCTTATATACCTCATTCAGATGATC 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
142 yAspSerThrAlaAlaAlaGlyAlaLeuLeuValIleGlnIleThr 158
148 TCCGAGCGCGCGAGATTCATCCCTTTTGAGGCGCTCCGCAATACAT 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
159 AlaGlnAlaAlaArgPheLeuTyrlleGlnGlnIleGlnIleArgAl 175
198 TAACACCGCGGAGCTCATTTCTTCCGACATGATGCTGCGAGGTGAGA 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
175 aTyArgAspGlu.....ValProSerLeuAlaThrIleSerLeuGlu 190
248 CTAGTTGGGGCCCAACATCCAGCAAGTCCAG.....CAGCTTACGGAT 291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 snSerTIPSerGlyLeuSerLysGlnIleGlnLeuAlaGlnIleAsn 206
292 GCGGTTTTTATATACCATTTGCGTTGGGTATATCCACCGGTAACTTGT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207 GlyIlePheArGlnThrProIleValIleValAlaAspAsnLysGlyAsnArgVa 223
342 GACGTGATGAGCAATGTCGCGACGTGATCCGACCTTAGCATGATGTGT 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 GlnIleThrAsnValThrSerLysValValThrSerAsnIleGlnLeuL 240
392 TYGTA 396
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
240 euLeu 241

```

seq\_name: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:US-10-127-890-6

```

seq_documentation_block:
; Sequence 6, Application US/10127890
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
;               Carroll, Stephen F.
;               Studnicka, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
;                   Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/127.890
; FILING DATE: 23-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994

```

```

; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MONICHOIAS, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-9155
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-127-890-6

```

alignment\_scores:  
Quality: 153.50 Length: 135  
Ratio: 1.651 Gaps: 5  
Percent Similarity: 68.889 Percent Identity: 32.593

alignment\_block:  
US-09-627-165b-15 x US-10-127-890-6 ..

Align seg 1/1 to: US-10-127-890-6 from: 1 to: 247

```

1 TACACAGATCTGAGCGATACGCCGCTCAT...AGGACACAGATCCCTCT 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
111 TyrgluarGleuglnThrAlaAlaGlyLysIleArgGlnsnIleProle 127
48 GGGTTATAGAGCAATTCATCCGCTCTGCGCGCTTCCTTATCCAGCGC 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127 uGlyLeuProAlaLeuAspSerAlaIleThrThrLeuPheTyrrAsnA 144
98 GCAGACCGCGCGCCAGCTCGTCCCTTATATATCCATTCATTCAGATGC 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
144 IAsnSerAlaAlaSerAla.....LeuMetValIleIleGlnSerThr 158
148 TCCGAGCGCGCGAGATTCATCCCATCTTTTGAGGCGCTCCGCAATACAT 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
159 SerGlnAlaAlaArgTyrrLysPheIleGlnGlnIleGlnLysArgVa 175
198 TAACACCGCGGAGTCATTTCTCCGACATGTACATGCTCGACGTCGAGA 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
175 IAsp.....LysThrPheLeuProSerLeuAlaIleIleSerLeuGlu 190
248 CTAGTTGGGCCCAACATCCAGCAAGTCCAG.....CAGCTTACGGAT 291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 snSerTIPSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsn 206
292 GCGGTTTTTATATACCATTTGCGTTGGGTATATCCACCGGTAACTTGT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207 GlyGlnPheGlnSerProValValIleuIleAsnAlaGlnsnIleVal 223
342 GACGTGAGCAATGTTCCGAC.....GTGATGCCAGCGTTCAGCATCA 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 aThrIleThrAsnValAlaAspAlaGlyValThrSerAsnIleAlaLeuL 240
386 TGTTG 390
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
240 euLeu 241

```

seq\_name: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:US-10-127-890-9



Wed Jul 31 08:46:44 2002

us-09-627-165b-15.rapn

Page 9

REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 102:  
US-10-127-890-102

alignment\_scores:  
Quality: 147.00 Length: 139  
Ratio: 1.690 Gaps: 3  
Percent Similarity: 62.590 Percent Identity: 31.655

alignment\_block:

US-09-627-165b-15 x US-10-127-890-102 ..

Align seg 1/1 to: US-10-127-890-102 from: 1 to: 251

```
1 TACACAGATCTGAGCAGTATCCGCTCATAGGAGCAGATCCCTCTGG 50
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
113 TyrProSerLeuGluGlyGlyLysAlaTyrArgGluThrAspLeuG 129
51 TATAGAGAACTC.....ATTCATCCGCTCGCGCTTCGTATC 91
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129 yllGluProLeuArgllGlyllLysLysLeuAspGluAsnAlaIle 146
92 CAGCGGAGCAGCAGCGGCGCCAGCGCTCCCTTATATCTCATTCAG 141
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146 spasnTyrLysProThrGluIleAlaSerSerLeuLeuValIleGln 162
142 ATGATCTCCGAGCGCGGAGATTCATCCATCTTTGGAGGCTCGGCA 191
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
163 MetValSerGluAlaIleArgPheThrPheIleGluAsnGlnIleArg 179
192 ATACATTAAACAGCGGAGGAGTATCTCCCGACATGATGCTGAGC 241
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
179 n.....AsnPheGlnGlnArgIleArgProAlaAsnAsnThrIleSer 194
242 TGGAGACTAGTGGGCGCCAAACATCCACGACATCCAGACAGCT...AGC 288
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
194 euGluAsnLysTyrGlyLysLeuSerPheGlnIleArgThrSerGlyAla 210
288 GATGCGCTTTTATATACCCATTGCGGTGGTATATCCACCGTAACTT 338
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
211 AsnGlyMetPheSerGluAlaValGluLeuGluArgAlaAsnGlyLys 227
339 CGTAGCTTGAGCATGTTCCGAGCTGATCCGACGCTTATGACATCATG 388
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
227 sTyrTyrValThrAlaValAspGlnValLysProLysIleAlaLeuLeu 244
389 TGTTCATGATGAGCAG 405
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
244 yspheValLysLysasp 249
```

seq\_name: /sgn2\_6/plodata/1/pae/US10\_NEW\_COMB.pep:US-10-127-890-111

seq\_documentation\_block:  
; Sequence 111, Application US/10127890  
; GENERAL INFORMATION:  
; APPLICANT: Belter, Marc D.  
; Caroll, Stephen F.  
; Studinka, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; Proteins  
; NUMBER OF SEQUENCES: 173  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Heid & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127, 890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 111:

SEQUENCE CHARACTERISTICS:

LENGTH: 251 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 111:

US-10-127-890-111

alignment\_scores:  
Quality: 147.00 Length: 139  
Ratio: 1.690 Gaps: 3  
Percent Similarity: 62.590 Percent Identity: 31.655

alignment\_block:

US-09-627-165b-15 x US-10-127-890-111 ..

Align seg 1/1 to: US-10-127-890-111 from: 1 to: 251

```
1 TACACAGATCTGAGCAGTATCCGCTCATAGGAGCAGATCCCTCTGG 50
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
113 TyrProSerLeuGluGlyGlyLysAlaTyrArgGluThrAspLeuG 129
51 TATAGAGAACTC.....ATTCATCCGCTCGCGCTTCGTATC 91
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129 yllGluProLeuArgllGlyllLysLysLeuAspGluAsnAlaIle 146
92 CAGCGGAGCAGCAGCGGCGCCAGCGCTCCCTTATATCTCATTCAG 141
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146 spasnTyrLysProThrGluIleAlaSerSerLeuLeuValIleGln 162
142 ATGATCTCCGAGCGCGGAGATTCATCCATCTTTGGAGGCTCGGCA 191
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
163 MetValSerGluAlaIleArgPheThrPheIleGluAsnGlnIleArg 179
192 ATACATTAAACAGCGGAGGAGTATCTCCCGACATGATGCTGAGC 241
```

```

179 n.....AsnpheGInGInaRgIleArGpRroAlaAsnaSnaThrIleSerL 194
      :      |||      :      :      :      :      :      :      :
242 TGGAGACTGTGGGGCCACACATCCACGACATCCAGCATCT...ACG 288
      |||:::|||||::: ||| |||::: ||| :
194 euGlunLysrIpGlYlYsLeuSerPheGInIleArGThrSerGlYAla 210
      :      :      :      :      :      :      :      :
289 GATGCGTTTAAATACCATTCCTGGTGTATATCCACCGTAACTT 338
      :      :      :      :      :      :      :      :
211 AsnGlYMetPheSerGlAlaValaGlutLeuGlAlaValaAsnGlYlYsLy 227
      :      :      :      :      :      :      :      :
339 CGTACGATGAGCATGTTCGCGAGCATGCGCAGCTTAAGCATCATGT 388
      :      :      :      :      :      :      :      :
227 sTYrTYrValThrAlaValaAspGlnValLYsProLYsIleAlaLeuL 244
      :      :      :      :      :      :      :      :
389 TGTGTGTATGTAGGAC 405
      |||:::|||||::: |||
244 yspheValcYsLYsAsp 249

```

seq\_name: /cgn2\_6/ptodata/1/paa/US09\_NFM\_COMB.pep:us-09-663-826-7

seq\_documentation\_block:

```

; Sequence 7, Application US/09663826
; GENERAL INFORMATION:
; APPLICANT: Russel, David R.
; FULLER, James T.
; TITLE OF INVENTION: METHOD FOR PRODUCING ANTIBODIES AND
; PROTEIN TOXINS IN PLANT CELLS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/663,826
; FILING DATE: 15-Sep-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,486
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 670513.90261
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-663-826-7

```

alignment\_scores:

```

Quality: 145.50      length: 135
Ratio: 1.565          Gaps: 5
Percent Similarity: 68.889      Percent Identity: 31.111

```

alignment\_block:

US-09-627-165b-15 x US-09-663-826-7 ..

Align seg 1/1 to: US-09-663-826-7 from: 1 to: 248

```

1 TACACGATCTGGAGCGATACGCCGTCAT...AGGACACGATCCCTCT 47
      |||:::|||||::: ||| |||::: ||| :
112 TYGluArGLeuGInThrAlaAlaGlyLYsIleArGluAsnIleProLe 128
      :      :      :      :      :      :      :      :
48 GGGTAAAGAGAAATCATTCATCCGCTCGGCGCTTGTATCCAGCG 97
      :      :      :      :      :      :      :      :
128 uGlLeuProAlaLeuAspSerAlaIleThrThrLeuTYrTYrTYrThra 145
      :      :      :      :      :      :      :      :
98 GCAGCACCCCGCCCAACGTCGTTCCCTTAATCTCTTACATGATC 147
      :      :      :      :      :      :      :      :
145 lAserSerAlaIleSerAla.....LeuLeuValLeuIleIleSerThr 159
      :      :      :      :      :      :      :      :
148 TCCGAGCGCCCGACATTCATCCCATCTTTGGAGGCGTCGCCATFACAT 197
      :      :      :      :      :      :      :      :
160 AlAGluSerAlaArGYrLYsPheIleGluGlnIleGlyLYsArGva 176
      :      :      :      :      :      :      :      :
198 TACACCGGGGAGTCATCTTCCCGACATGTACATGTCGATGAGAGA 247
      :      :      :      :      :      :      :      :
176 lAsp.....LYsThrPheLeuProSerLeuAlaThrIleSerLeuGluA 191
      :      :      :      :      :      :      :      :
248 CTAGTGGGCGCCACATCCACGACATCCAG.....CAGTCTACGAT 291
      :      :      :      :      :      :      :      :
191 snAsnTrpSerAlaLeuSerLYsGlnIleGlnIleAlaSerThrAsn 207
      :      :      :      :      :      :      :      :
292 GCGCTTTTAATACCATTCGCTTGGGTATATCCACCGCTAACTTGT 341
      :      :      :      :      :      :      :      :
208 GLYGlNpHeGluSerProValaIleuIleAspLYsAsnGlnArGva 224
      :      :      :      :      :      :      :      :
342 GAGCTTGAGCAAT.....GTTCGCGAGCATCGCACCTTACGATCA 385
      :      :      :      :      :      :      :      :
224 lSerIleThrAsnAlaSerAlaArGValaIleThrSerAlaIleAlaLeuL 241
      :      :      :      :      :      :      :      :
386 TGTGT 390
      :      :      :      :      :      :      :      :
241 leuL 242

```

seq\_name: /cgn2\_6/ptodata/1/paa/US09\_NFM\_COMB.pep:us-09-663-826-15

seq\_documentation\_block:

```

; Sequence 15, Application US/09663826
; GENERAL INFORMATION:
; APPLICANT: Russel, David R.
; FULLER, James T.
; TITLE OF INVENTION: METHOD FOR PRODUCING ANTIBODIES AND
; PROTEIN TOXINS IN PLANT CELLS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/663,826
; FILING DATE: 15-Sep-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,486
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 670513.90261
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000

```





```
187 ||:::||||: || ||::: ||:::
    luAsnGlnTrpSerAlaLeuSerLysGlnIlePheLeuAlaGlnAsnGln 203
289 GATGGCGTTTAAATACCAATTTCGTTGGGTATATCCACCGTAACTT 338
    || ||:::||||: || ||:::||||:
204 GlyGlyLysPheArgAsnProValAspLeuIleLysProThrGlyGluAr 220
339 CGTGACGTTGAGCAATGTTCCGACGTGATCGCCAGCTTACGATCATGT 388
    :::::||||: :::::||||: || |
220 gPheGlnValThrAsnValAspSerAspValValLysGlyAsnIleLysL 237
389 TGTGTATGTAGGACGACCATCTTCTCCGACGTCCCTAT 432
    ||:::||||: ||:::||||: ||:::
237 euLeuLeu...AsnSerArgAlaSerThrAlaAspGluAsnPhe 250
```